Abstract
We provide general results on the consistency of certain bootstrap methods applied to degree-2 degenerate statistics of $U$- and $V$-type. While it follows from well known results that the original statistic converges in distribution to a weighted sum of centered 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.

2000 Mathematics Subject Classification. Primary 62G09, 62F40; secondary 62G20, 62F05.
Keywords and Phrases. Bootstrap, consistency, $U$-statistics, $V$-statistics.
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 by Bickel and Freedman (1981) for Efron’s bootstrap applied to linear and related statistics, Arcones and Giné (1992) for Efron’s bootstrap in connection with $U$- and $V$-statistics, and Stute, González Manteiga and Presedo Quindimil (1993) 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 $U$- and $V$-type. Statistics of this type often emerge from goodness-of-fit tests; see for example de Wet (1987, Sect. 2). Under usual assumptions, the limit distribution of such a $U$-statistic will be that of a weighted sum of independent and centered $\chi^2$ random variables with 1 degree of freedom. In a few cases this distribution is actually known; see for example Darling (1955, Sect. 7) and Moore and Spruill (1975). 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 (1981), for proving consistency of Efron’s bootstrap for statistics as the sample mean or related quantities, and modified by Dehling and Mikosch (1994), 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; 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 (1994) 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 (1998) and Jiménez-Gamero, Muñoz-García and Pino-Mejías (2003). In both cases the limit distribution of the original statistic was derived via a spectral decomposition of the kernel. Jiménez-Gamero et al. (2003) mimicked this proof also on the bootstrap side while Fan (1998) dismissed this possibility and employed empirical process arguments. While Jiménez-Gamero et al. (2003) and Fan (1998) 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 Stute et al. (1993). 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 $(\Omega, A, P)$, with common distribution function $F_\theta$, where $\theta \in \Theta \subseteq \mathbb{R}^p$.

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

(iii) $0 < \int h^2(x, y; \theta) \, dF_\theta(x) \, dF_\theta(y) < \infty$.

It is well known (see, for example, Serfling (1980, p. 194) or Lee (1990, Theorem 3.2.2.1, p. 90) that under $(A1)$ the following result holds true (‘$d \to$’ denotes convergence in distribution):

\[ U_n = \frac{1}{n} \sum_{j=1}^{n} \sum_{k \neq j} h(X_{nj}, X_{nk}; \theta) \quad d \to Z := \sum_{\nu=1}^{\infty} \lambda_\nu (Z_\nu^2 - 1), \tag{2.1} \]

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

\[ \int h(x, y; \theta) g(y) \, dF_\theta(y) = \lambda g(x). \]

Furthermore, since $\sum_{\nu=1}^{\infty} \lambda_\nu^2 = Eh^2(X_1, X_2; \theta) < \infty$ (see Serfling (1980, 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 that $X_{n1}, \ldots, X_{nn}$ are independent with a distribution converging to that of $X_1$. We make the following assumption:
(A2)(i) $(X_{nj})_{j=1,...,n}, n \in \mathbb{N}$, is a triangular scheme of random variables defined on respective probability spaces $(\Omega_n, A_n, P_n)$, where $X_{n1}, \ldots, X_{nn}$ are independent with common distribution function $F_n$. Furthermore, it holds that $F_n \xrightarrow{d} F_\theta$ and $\hat{\theta}_n \xrightarrow{n\to\infty} \theta$. (‘$\xrightarrow{d}$’ denotes convergence in distribution.)

(ii) $h : \mathbb{R}^d \times \mathbb{R}^d \times \mathbb{R}^p \to \mathbb{R}$ is a measurable function, the set $\delta h$ of its discontinuity points fulfills $\int I((x, y, \theta) \in \delta h) \ dF_\theta(x) \ dF_\theta(y) = 0$, and $\int h(x, y; \hat{\theta}_n) \ dF_n(x) = 0 \ \forall y \in \mathbb{R}^d$, that is, the kernel $h(\cdot, \cdot; \hat{\theta}_n)$ is degenerate under $F_n$.

(iii) $Eh^2(X_{n1}, X_{n2}; \hat{\theta}_n) \xrightarrow{n\to\infty} Eh^2(X_1, X_2; \theta)$.

Remark 1. Under (A2)(i) the assumption (A2)(iii) is equivalent to the uniform integrability of $(h^2(X_{n1}, X_{n2}; \hat{\theta}_n))_{n \in \mathbb{N}}$, what 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 $n \to \infty$,

$$U_{nn} \xrightarrow{d} Z,$$

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

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

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 $\hat{F}_n$ of the unknown distribution function $F_\theta$. A minimal property that one usually expects for such a resampling scheme is that $\hat{F}_n \xrightarrow{d} F_\theta$, in probability or almost surely. Furthermore, we also assume that $\hat{\theta}_n$ is a consistent estimator of $\theta$. (In the case of a model-based bootstrap method the $X^*_j$ will be typically drawn from $F_{\hat{\theta}_n}$; however, we do not require this here.) The hope is that $U_n$ is now consistently mimicked by

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

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 $X^*_1 \xrightarrow{d} X_1$ in probability. Moreover, $\hat{\theta}_n \xrightarrow{p} \theta$. (‘$\xrightarrow{p}$’ denotes convergence in probability.)

(ii) $h : \mathbb{R}^d \times \mathbb{R}^d \times \mathbb{R}^p \to \mathbb{R}$ is a measurable function, the set $\delta h$ of its discontinuity points fulfills $\int I((x, y, \theta) \in \delta h) \ dF_\theta(x) \ dF_\theta(y) = 0$, and $E(h(X^*_1, y; \hat{\theta}_n) \mid X_1, \ldots, X_n) = 0 \ \forall y \in \mathbb{R}^d$, that is, the kernel $h(\cdot, \cdot; \hat{\theta}_n)$ is degenerate for $X^*_1$.

(iii) $E(h^2(X^*_1, X^*_2; \hat{\theta}_n) \mid X_1, \ldots, X_n) \xrightarrow{p} Eh^2(X_1, X_2; \theta)$.

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

$$U^*_n \xrightarrow{d} Z \quad \text{in probability},$$
where $Z$ is defined in (2.1) above. Moreover,

$$
\sup_{-\infty < x < \infty} |P(U_n \leq x \mid X_1, \ldots, X_n) - P(U_n \leq x)| \overset{P}{\to} 0.
$$

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 2. Jiménez-Gamero et al. (2003) derived a similar result by a different method of proof. They studied the special case where the kernel has the form

$$
h(x, y; \theta) = \int q(x, t; \theta)q(y, t; \theta)\,dG_\theta(t),
$$

for some function $q$ and some finite measure $G_\theta$. Under smoothness conditions on $q$, $G_\theta$ and the densities of the random variables involved, they showed that the eigenvalues of some operators connected with $h(X^*_j, X^*_k; \hat{\theta}_n)$ 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 (1977, p. 115) mentioned a few cases where it might be preferable to use $U$- 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

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

which implies by (2.1) that

$$
V_n \overset{d}{\to} Z + Eh(X_1, X_1; \theta). \quad (2.2)
$$

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

$$
V_{nn} = \frac{1}{n} \sum_{i,j=1}^{n} h(X_{nj}, X_{nk}; \hat{\theta}_n).
$$

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

(A4) The set $\delta h$ of discontinuity points of $h$ fulfills $\int I((x, x, \theta) \in \delta h)\,dF_\theta(x) = 0$ and it holds that $E h(X_{n1}, X_{n1}; \hat{\theta}_n) \to_{n \to \infty} Eh(X_1, X_1; \theta)$. 
Lemma 2.2. Suppose that (A1) to (A4) are fulfilled. Then, as $n \to \infty$,
\[ V_{nn} \xrightarrow{d} Z + Eh(X_1, X_1; \theta) \]
and
\[ \sup_{-\infty < x < \infty} |P_n(V_{nn} \leq x) - P(V_n \leq x)| \xrightarrow{n \to \infty} 0. \]

Now we can easily identify sufficient conditions for $V^*_n$ converging to the same limit as $V_n$. We will assume that (A4*)
\[ \text{The set } \delta_h \text{ of discontinuity points of } h \text{ fulfills } \int I((x, x, \theta) \in \delta_h) \, dF(\theta(x)) = 0 \]
and it holds that $E \left( h(X^*_1, X^*_1, \hat{\theta}_n) \mid X_1, \ldots, X_n \right) \overset{P}{\to} Eh(X_1, X_1; \theta)$.

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 $n \to \infty$,
\[ V^*_n \xrightarrow{d} Z + Eh(X_1, X_1; \theta) \text{ in probability,} \]
where $Z$ is defined in (2.1) above. Moreover,
\[ \sup_{-\infty < x < \infty} \left| P(V^*_n \leq x \mid X_1, \ldots, X_n) - P(V_n \leq x) \right| \overset{P}{\to} 0. \]

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 Lemma 2.2.

3. A GOODNESS-OF-FIT TEST FOR THE NIG-MODEL

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 very critical. Barndorff-Nielsen (1997) proposed to model the process of logarithmic asset prices $(Y_t)_{t \geq 0}$ by a normal inverse Gaussian (NIG hereafter) process of Lévy type. A NIG process with parameters $\alpha, \beta, \mu$ and $\delta$ is a particular Lévy process where $Y(t)$ has a normal inverse Gaussian density $f_{\text{NIG}}(\cdot; \alpha, \beta, \mu_t, \delta_t)$,
\[ f_{\text{NIG}}(x; \alpha, \beta, \mu_t, \delta_t) = \frac{\alpha \delta t}{\pi} \exp \left( \delta t \sqrt{\alpha^2 - \beta^2} + \beta(x - \mu_t) \right) \frac{K_1(\alpha \sqrt{(\delta t)^2 + (x - \mu_t)^2})}{\sqrt{(\delta t)^2 + (x - \mu_t)^2}}. \]

Here $\alpha, \beta > 0$, $\beta \in (-\alpha, \alpha)$, $\mu \in \mathbb{R}$, and $K_1$ denotes the modified Bessel function of third order and index 1, that is, $K_1(z) = (1/2) \int_0^\infty \exp(-z(u + (1/u))/2) \, du$. 
We assume that we observe the asset prices at equidistant time points $\Delta, 2\Delta, \ldots, n\Delta$ 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 $X_j = Y(j) - Y(j - 1)$ $(j = 1, \ldots, n)$ are independent and have a NIG distribution with parameters $\alpha$, $\beta$, $\mu$ and $\delta$. While their common density is rather complicated their characteristic function has the following simple closed form (see equation (3.7) in Barndorff-Nielsen (1997)):

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

Since empirical studies support distributions which are negative skewed and have positive mean we choose the parameter space (see also the characteristics of the NIG distribution below)

$$\Theta = \left\{ \theta = (\alpha, \beta, \mu, \delta) \mid \alpha > -\beta > 0, \delta > 0, \mu > -\frac{\delta \beta}{\sqrt{\alpha^2 - \beta^2}} \right\}.$$ 

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

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

We consider the following test statistic:

$$T_n = n \int_{-\infty}^{\infty} \left| \hat{c}_n(t) - c(t; \hat{\theta}_n) \right|^2 w(t) \, dt,$$

where $\hat{c}_n(t) = n^{-1} \sum_{j=1}^{n} \exp(itX_j)$ is the empirical characteristic function of the increments, $\hat{\theta}_n$ is some estimator of $\theta$ and $w : \mathbb{R} \rightarrow [0, \infty)$ 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 $\int (1 + |t|)^{4}w(t) \, dt < \infty$ and vanishes only on a set of Lebesgue measure zero.

Test statistics of this type have also been considered by Fan (1998) where the consistency of bootstrap approximations is proven under a set of conditions different to ours.

For definiteness, we restrict our attention to a method-of-moments estimator $\hat{\theta}_n$ of $\theta$. Under the null hypothesis, the distribution of the increments $X_j$ has the following characteristics (see, for example Schoutens (2005)):

- mean: $\mu + \frac{\delta \beta}{\alpha^2 - \beta^2}$,
- variance: $\frac{\alpha^2 \delta^{3/2}}{(\alpha^2 - \beta^2)^{1/2}}$,
- skewness: $\frac{3\beta}{\alpha \delta^{1/2} (\alpha^2 - \beta^2)^{1/4}}$,
- excess: $3 \left( 1 + \frac{\alpha^2 + 4\beta^2}{\delta \alpha^2 \sqrt{\alpha^2 - \beta^2}} \right)$.

Accordingly, we obtain $\hat{\theta}_n = (\hat{\alpha}_n, \hat{\beta}_n, \hat{\mu}_n, \hat{\delta}_n)'$ by equating the first four theoretical moments with their empirical counterparts. It can be shown that $\theta$ is a twice differentiable function of the first four theoretical moments and vice versa that these
moments are continuous functions of \( \theta \). Similar results are obtained for \( \hat{\theta}_n \) applying the empirical moments instead. Thus, by Taylor expansion, there exists some function \( g : \mathbb{R}^4 \times \mathbb{R} \rightarrow \mathbb{R}^4 \) such that

\[
\hat{\theta}_n - \theta = \frac{1}{n} \sum_{j=1}^{n} g(\theta; X_j) + o_P(n^{-1/2}) \tag{3.2}
\]

and \( Eg(\theta; X_j) = 0, E\|g(\theta; X_j)\|_2^2 < \infty \), where \( \|x\|_2 = \sqrt{\sum_i x_i^2} \). This implies in particular that

\[
\hat{\theta}_n - \theta = O_P(n^{-1/2}). \tag{3.3}
\]

Since \( c \) is twice differentiable with respect to \( \theta \) a Taylor series expansion leads to

\[
T_n = n \int_{-\infty}^{\infty} \left| \hat{c}_n(t) - c(t; \theta) - Dc(t; \theta)(\hat{\theta}_n - \theta) - \frac{1}{2}(\hat{\theta}_n - \theta)'D^2c(t; \theta)(\hat{\theta}_n - \theta) \right|^2 w(t) \, dt,
\]

for some \( \hat{\theta}(t) \) between \( \hat{\theta}_n \) and \( \theta \), where \( Dc \) and \( D^2c \) denote the gradient (row) vector and the Hessian matrix of \( c \) with respect to \( \theta \), respectively. Since assumption (A5) ensures the \( w \)-integrability of \( \|D^2c(t; \theta(t))\|^2 \) \( (\|A\|^2 = \lambda_{\text{max}}(A'A)) \) we obtain by (3.3) that

\[
T_n = I_n + o_P(1),
\]

where \( I_n = n \int_{-\infty}^{\infty} \left| \hat{c}_n(t) - c(t; \theta) - Dc(t; \theta)(\hat{\theta}_n - \theta) \right|^2 w(t) \, dt \). Furthermore, (3.2) allows us to approximate further:

\[
T_n = I_n^0 + o_P(1),
\]

where \( I_n^0 = \int_{-\infty}^{\infty} \left| n^{-1/2} \sum_{j=1}^{n} \{ \exp(itX_j) - c(t; \theta) - Dc(t; \theta)g(\theta; X_j) \} \right|^2 w(t) \, dt \). The statistic \( I_n^0 \) is a degenerate \( V \)-statistic with a symmetric kernel

\[
h(x, y; \theta) = \int_{-\infty}^{\infty} \{ \exp(itx) - c(t; \theta) - Dc(t; \theta)g(\theta; x) \} \times \{ \exp(-ity) - c(-t; \theta) - Dc(-t; \theta)g(\theta; y) \} w(t) \, dt.
\]

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 \( \theta \), 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

\[
T_n \xrightarrow{d} \sum_{\nu=1}^{d} \lambda_\nu (Z^2_\nu - 1) + Eh(X_1, X_1; \theta),
\]

where \( Z_1, Z_2, \ldots \) are independent standard normal random variables and the \( \lambda_\nu \) are the eigenvalues of the equation

\[
E [h(x, X_1; \theta)g(X_1)] = \lambda g(x).
\]
To implement a test which has asymptotically a prescribed size \( \alpha \), 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 \( \lambda_n \) which in turn depend on the true parameter \( \theta \) in a complicated way. Therefore, we propose the following bootstrap procedure:

(1) Given \( X_1, \ldots, X_n \), estimate \( \theta \) by \( \hat{\theta}_n \) and generate a sample \( X_1^*, \ldots, X_n^* \) with a common density \( f_{NIG}(\cdot; \hat{\theta}_n) \). According to Barndorff-Nielsen (1997, p. 2), a random variable \( X \) with density \( f_{NIG}(\cdot; \alpha, \beta, \mu, \delta) \) can be generated by drawing first a random variable \( Z \) with an inverse Gaussian distribution with parameters \( \delta \) and \( \sqrt{\alpha^2 - \beta^2} \), and then, conditioned on \( Z = z \), drawing \( X \sim N(\mu + \beta z, z) \). See also Schoutens (2005, p. 111 f.) for simulating an inverse Gaussian distributed random variable. The bootstrap sample conditioned on \( X_1, \ldots, X_n \) satisfies \( H_0 \) with \( \hat{\theta}_n \) instead of \( \theta \).

(2) Define a bootstrap counterpart \( \hat{\theta}_n^* \) to \( \hat{\theta}_n \) which is based on the bootstrap sample by the method of moments and set \( \hat{c}^*_n(t) = n^{-1} \sum_{j=1}^n \exp(itX^*_j) \). Then

\[
T^*_n = n \int_{-\infty}^{\infty} \left| \hat{c}^*_n(t) - c(t; \hat{\theta}_n^*) \right|^2 w(t) \, dt
\]

is the bootstrap version of our test statistic \( T_n \).

(3) Define the critical value \( t^*_n \) as the \( (1 - \alpha) \)-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 \( (1 - \alpha) \)-quantile of the empirical distribution associated with \( T^*_n, \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 (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

\[
\hat{\theta}_n \xrightarrow{P-a.s.} \theta. \tag{3.4}
\]

Analogously to (3.2), we obtain that

\[
\hat{\theta}_n^* - \hat{\theta}_n = \frac{1}{n} \sum_{j=1}^n g(\hat{\theta}_n; X_j^*) + R_n^*, \tag{3.5}
\]

where

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

(3.4) and (3.5) yield that

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

where \( I_n^{0*} = n^{-1} \sum_{j,k=1}^n h(X^*_j, X^*_k; \hat{\theta}_n) \). It follows from the construction that

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

And finally, we obtain from (3.4) that

\[
E(h^2(X_1^*, X_2^*; \hat{\theta}_n) \mid X_1, \ldots, X_n) \xrightarrow{P-a.s.} Eh^2(X_1, X_2; \theta)
\]

and

\[
E(h(X_1^*, X_1^*; \hat{\theta}_n) \mid X_1, \ldots, X_n) \xrightarrow{P-a.s.} Eh(X_1, X_1; \theta).
\]

To summarize we have verified that the assumption (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
\[ \sup_{-\infty < x < \infty} |P(T_n^* \leq x \mid X_1, \ldots, X_n) - P(T_n \leq x)| \xrightarrow{P-a.s.} 0. \]
For \( t_\alpha = \inf\{t : P(T_n^* \leq t \mid X_1, \ldots, X_n) \geq 1 - \alpha\} \), we obtain that
\[ P(T_n > t_\alpha) \xrightarrow{n \to \infty} \alpha, \]
that is, the test has asymptotically the correct size.

4. Proofs

The main idea of the proof of Lemma 2.1 is similar to that of Theorem 2.1 in Dehling and Mikosch (1994) 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 \( \theta \). Hence, the proof requires some modifications. For 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 Billingsley (1999, 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 \( g : \Omega_0 \to \mathbb{R}^d \) and \( g_n : \Omega_0 \to \mathbb{R}^d \) with \( \widetilde{X}_j := g(\omega_j) \sim F_\theta \), \( \widetilde{X}_{nj} := g_n(\omega_j) \sim F_n \) and
\[ \widetilde{X}_{nj} - \widetilde{X}_j \xrightarrow{P-a.s.} 0, \quad (4.1) \]
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 \( \theta_n \xrightarrow{n \to \infty} \theta \) and the assumed continuity property of \( h \), that, for \( 1 \leq j, k \leq n \),
\[ h(\widetilde{X}_{nj}, \widetilde{X}_{nk}; \theta_n) = h(\widetilde{X}_j, \widetilde{X}_k; \theta) \xrightarrow{P-a.s.} 0. \quad (4.2) \]
It follows from \( Eh^2(\widetilde{X}_{n1}, \widetilde{X}_{n2}; \theta_n) \xrightarrow{n \to \infty} Eh^2(\widetilde{X}_1, \widetilde{X}_2; \theta) \) 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.
This, however, implies, in conjunction with the strong law of large numbers, that
\[ E_P \left( h(\tilde{X}_{n1}, \tilde{X}_{n2}; \theta_n) - h(\tilde{X}_1, \tilde{X}_2; \theta) \right)^2 \xrightarrow{n \to \infty} 0. \] (4.3)

Observe now that
\[ \bar{U}_n - \bar{U}_m = \frac{1}{n} \sum_{j=1}^{n} \sum_{k \neq j} k_n((\tilde{X}_j, \tilde{X}_n), (\tilde{X}_k, \tilde{X}_n)) \]
is a degenerate U-statistic of the independent random variables \((\tilde{X}_1, \tilde{X}_n), \ldots, (\tilde{X}_n, \tilde{X}_n)\)
and with kernel \(k_n((x, x'), (y, y')) = h(x, y; \theta) - h(x', y'; \theta_n)\). We can easily compute that
\[ E_P \left( \bar{U}_n - \bar{U}_m \right)^2 = \frac{2(n-1)}{n} E_P \left( k_n((\tilde{X}_1, \tilde{X}_n1), (\tilde{X}_2, \tilde{X}_n2)) \right)^2 \xrightarrow{n \to \infty} 0, \]
which immediately implies the first assertion of this lemma. Furthermore, the second assertion follows since the limit distribution of \(U_n\) is continuous. \(\square\)

**Proof of Lemma 2.2**. Here we employ exactly the same coupling as in the proof of Lemma 2.1. We denote by \(\bar{V}_n\) and \(\bar{V}_{nn}\) the copies of \(V_n\) and \(V_{nn}\) on the probability space \((\Omega, \mathcal{A}, \bar{P})\), respectively. We have that
\[ \bar{V}_{nn} = \bar{U}_{nn} + \frac{1}{n} \sum_{j=1}^{n} h(\tilde{X}_{nj}, \tilde{X}_{nk}; \theta_n). \]
We obtain from \(E(h(\tilde{X}_{n1}, \tilde{X}_{n1}; \theta_n) \xrightarrow{n \to \infty} E(h(\tilde{X}_1, \tilde{X}_1; \theta)) and (4.2)\) that \((h(\tilde{X}_{n1}, \tilde{X}_{n1}; \theta_n))_{n \in \mathbb{N}}\) is a uniformly integrable family of random variables. Hence, it follows from (4.2) that
\[ E_{\bar{P}} \left| h(\tilde{X}_{n1}, \tilde{X}_{n1}; \theta_n) - h(\tilde{X}_1, \tilde{X}_1; \theta) \right| \xrightarrow{n \to \infty} 0. \]
This, however, implies, in conjunction with the strong law of large numbers, that
\[ E_{\bar{P}} \left| \frac{1}{n} \sum_{j=1}^{n} h(\tilde{X}_{nj}, \tilde{X}_{nj}; \theta_n) - E_P h(X_1, X_1; \theta) \right| \]
\[ \leq E_{\bar{P}} \left| h(\tilde{X}_{n1}, \tilde{X}_{n1}; \theta_n) - h(\tilde{X}_1, \tilde{X}_1; \theta) \right| + E_{\bar{P}} \left| \frac{1}{n} \sum_{j=1}^{n} h(\tilde{X}_j, \tilde{X}_j; \theta) - E_P h(X_1, X_1; \theta) \right| \]
\[ \xrightarrow{n \to \infty} 0. \]
Therefore, we obtain that
\[ \bar{V}_{nn} = \bar{U}_{nn} + E(h(X_1, X_1; \theta)) + o_P(1), \]
which yields, in conjunction with Lemma 2.1, the assertions of the lemma. \(\square\)

**References**


