Scientific
Research
Publishing

High Moments Jarque-Bera Tests for Arbitrary Distribution Functions

Gane Samb Lo[1,2], Oumar Thiam[2], Mohamed Cheikh Haidara[2]

[1]Laboratoire de Statistique Théorique et Appliquée (LSTA), Université Pierre et Marie Curie, Paris, France
[2]LERSTAD, Université Gaston Berger de Saint-Louis, Saint-Louis, Senegal
Email: gane-samb.lo@ugb.edu.sn, ganesamblo@ganesamblo.net, othiam@univi.net, mhaidara@univi.net

Abstract

The Jarque-Bera's fitting test for normality is a celebrated and powerful one. In this paper, we consider general Jarque-Bera tests for any distribution function (df) having at least $4k$ finite moments for $k \geq 2$. The tests use as many moments as possible whereas the JB classical test is supposed to test only skewness and kurtosis for normal variates. But our results unveil the relations between the coeffients in the JB classical test and the moments, showing that it really depends on the first eight moments. This is a new explanation for the powerfulness of such tests. General Chi-square tests for an arbitrary model, not only normal, are also derived. We make use of the modern functional empirical processes approach that makes it easier to handle statistics based on the high moments and allows the generalization of the JB test both in the number of involved moments and in the underlying distribution. Simulation studies are provided and comparison cases with the Kolmogorov-Smirnov's tests and the classical JB test are given.

Keywords

Asymptotic Distribution, Asymptotic Statistical Tests, Normality Tests, Functional Empirical Processes

1. Introduction

In this paper, we are concerned with generalizations of Jarque-Bera's (JB) [1] tests based on arbitrary first ($4k$) moments, $k \geq 2$, rather than on the first eight ones as usual. (See [2] for a reminder of JB tests, page 69). We obtain general statistics that allow statistical tests for any distribution function G provided it has enough moments. For a reminder, the classical JB test belongs to the class of omnibus moment tests, *i.e.* those which assess simultaneously whether the skewness and kurtosis of the data are consistent with a Gaussian model. This test proves

optimum asymptotic power and good finite sample properties (see [1]). A detailed description of that test and related indepth analyses can be found in Bowman and Shenton, D'Agosto, D'Agostino *et al.*, etc. (see [3]-[5] and [6]).

Let $X, X_1, X_2, \cdots$ be a sequence of independent and identically distributed random variables (*r.v.*'s) defined on the same probability space $(\Omega, \mathcal{A}, \mathbb{P})$. For each $n \geq 1$, the skewness and kurtosis coefficients related to the sample $X, \cdots, X_n$ are defined by.

$$b_{n,2} = \frac{(1/n)\sum_{i=1}^{n}(X_1 - \bar{X})^3}{\left[(1/n)\sum_{i=1}^{n}(X_1 - \bar{X})^2\right]^{3/2}}, \quad a_{n,2} = \frac{(1/n)\sum_{i=1}^{n}(X_1 - \bar{X})^4}{\left[(1/n)\sum_{i=1}^{n}(X_1 - \bar{X})^2\right]^2} \tag{1}$$

These statistics are designed to estimate the theoretical skewness and kurtosis given by $b_2 = \mathbb{E}(X - m)^3/\sigma^3$ and $a_2 = \mathbb{E}(X - m)^4/\sigma^4$ where $m = \mathbb{E}(X)$ and $\sigma^2 = \text{var}(X)$ respectively denote the mean and the variance of X that is supposed to be nondegenerated. Here and in all the sequel, $\mathbb{E}$ stands for the mathematical expectation with respect to the probability $\mathbb{P}$. Now, under the hypothesis:

*H*0: X follows a Gaussian normal law, we have $b^2 = 0$ and $a = 3$ and the *JB* statistic

$$T_n = \frac{n}{6}\left(b_{n,2}^2 + \frac{1}{4}(a_{n,2} - 3)^2\right) \tag{2}$$

has an asymptotic chi-square distribution with two degrees of freedom under the null hypothesis of normality. Jarque-Bera's test consists in rejecting *H*0 when T_n is far from zero. We will find below that the constants 6 and 24 used in (2), actually, are closely related to the first four even moments of a $\mathcal{N}(0,1)$ random variable which are 1, 3, 15 and 105 and a more convenient form of (2) is

$$T_n = n\left(b_{n,2}^2/6 + (a_{n,2} - 3)^2/24\right)$$

Our objective here is to generalize *JB*'s test to a general *df G* by considering high moments $m_\ell = \mathbb{E}(X^\ell)$, $\ell \geq 1$, with $m_1 \equiv m$, instead of the first eight moments only. We base our methods on the remark that for a random variable $X \sim \mathcal{N}(m, \sigma^2)$, one has

$$\forall k \geq 0, \ \mathbb{E}\left((X - m)^{2k+1}\right) = 0, \ \mathbb{E}\left((X - m)^{2k}\right) = \frac{(2k)!}{2^k k!}\sigma^{2k}. \tag{H1}$$

Actually, *JB*'s test only checks the third and fourth moments of *X* while the coefficients of the *JB* statistic (2) uses the first eight moments of *X*. Our guess is that we would have better tests if we are able to simultaneously check all the first $(2k)$ moments for some $k \geq 2$. To this purpose, we consider the following statistics, that is the normalized centered empirical moments (NCEM),

$$b_{n,p} = \frac{\mu_{n,2p-1}}{\mu_{n,2}^{(2p-1)/2}} \quad \text{and} \quad a_{n,p} = \frac{\mu_{n,2p}}{\mu_{n,2}^p}, \quad p \geq 2 \tag{3}$$

where

$$m_{n,\ell} = \sum_{i=1}^{n} X_i^\ell \quad \text{and} \quad \mu_{n,\ell} = \frac{1}{n}\sum_{i=1}^{n}(X_i - \bar{X})^\ell, \ \ell \geq 1$$

are the ℓth non-centered and the centered empirical moments. By the classical law of large numbers, the statistics in (3) are, for each fixed *p*, asymptotic estimators of

$$b_p = \frac{\mathbb{E}\left((X - m)^{2p-1}\right)}{\sigma^{(2p-1)}} \quad \text{and} \quad a_p = \frac{E\left((X - m)^{2p}\right)}{\sigma^{2p}}, \quad p \geq 2, \tag{4}$$

whenever the $(4p)^{\text{th}}$ moment exists. Finally we consider C^1-class functions $(f_p)_{p \leq i \leq k}$ et $(g_p)_{1 \leq p \leq k}$ and denote $f = (f_1, \cdots, f_k)$ and $g = (g_1, \cdots, g_k)$.

Our general test is based on the following statistics, for $k \geq 2$,

$$T_n(f,g,k) = \sum_{p=2}^{k} \left(f_p\left(b_{n,p}\right) + g_p\left(a_{n,p}\right) \right) \tag{5}$$

which almost-surely $(a.s)$ tends to

$$T(f,g,k) = \sum_{p=2}^{k} \left(f_p\left(b_p\right) + g_p\left(a_p\right) \right), \tag{6}$$

as $n \to +\infty$. For an independent and identically distributed sequence $X_1, X_2, \cdots$ of *r.v.*'s associated with a distribution function *G* having a finite 2*k*-moment, we will have by Theorem 1 below that

$$T_n(f,g,k) - T(f,g,k) \xrightarrow{\;\mathbb{P}\;} 0 \quad \text{as} \quad n \to +\infty$$

From such a general result, we are able to derive a normality test by using it with $b_p = 0$, $a_p = (2p)/(2^p\, p!)$ for $2 \le p \le k$, and rejects normality for a large value of $T_n(f,q,k)$.

We are going to establish a general asymptotic normality of $T_n(f,q,k)$ for any *df*'s *G* with 4*k* finite moments. These results provide themselves efficient tests for an arbitrary *d.f.* Next, we will derive chi-square tests that generalize JB's test for higher moments and for arbitrary *df*'s too.

Our results will show that these tests based on the 2*k* moments, need, in fact, the eight 4*k* moments for computing the variance. This unveils that the classical JB's test is not based only on the kurtosis and the skewness but also on the sixth and the eighth moments. To describe the complete form of the Jarque-Bera method, put

$$aj(p) = \sigma^{-(4p)} E\left(X^{2p} - pE\left(X^{2p}\right)X^2 \right)^2 \quad \text{and} \quad bj(p) = \sigma^{-(4p-2)} E\left(X^{4p-2} \right)$$

The JB's test for a $\mathcal{N}(m,\sigma^2)$ *r.v.* will be showed to derive from the following general law

$$n\left(\left(b_{n,2} - b_p\right)^2 \big/ bj(p) + \left(a_{n,2} - a_p\right)^2 \big/ aj(p) \right) \sim \chi_2^2. \tag{7}$$

with the particular coefficients $p = 2$, $b_p = 0$ and $a_p = 3$. This may be a new explanation of the powerfulness of the JB classical tests since a successful test of normality means that the sample is from a *df* having same first eight moments as the $\mathcal{N}(0,1)$ *r.v.*, and this is very highly improbable for a non normal *r.v.*.

As an illustration of what proceeds, consider a distribution following a double-gamma distribution $\gamma_d\left(\left(1+\sqrt{13}\right)\big/2, 1\right)$ of density probability $f(x) = b^a\big/ 2\Gamma(a)|x|^{a-1}\exp\left(-b|x|\right)$ with $a = \left(1+\sqrt{13}\right)\big/2$. This *rv* is centered and has a kurtosis coefficient equal to 3. It is rejected from normality by the JB test. If only the skewned and kurtosis do matter, it would not be the case. Actually, the rejection comes from the parameters $aj(2)$ and $bj(2)$ that are very different from a standard normal distribution to this specific distribution.

The rest of the paper is organized as follows. In Subsection 2.1 of Section 2, we begin to give a concise of reminder the modern theory of functional empirical processes that is the main theoretical tool we use for finding the asymptotic law of (5). Next in Subsection 2.2, we establish general results of the consistency of (5) and its asymptotic law, consider particular cases in Subsection 2.3, propose chi-square universal tests in Subsection 2.4 and finally state the proofs in Subsection 2.5. We end the paper by Section 3 where simulation results concerning the normal and double-exponential models are given.

We here express that in all the sequel the limits are meant as $n \to +\infty$ and this will not be written again unless it is necessary.

2. Results and Proofs

2.1. A Reminder of Functional Empirical Process

Since the empirical functional process is our key tool here, we are going to make a brief reminder on this process associated with $X_1, X_2, \cdots$, and defined for each $n \ge 1$ by

$$\mathbb{G}_n(f) = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \left(f(X_i) - \mathbb{E}f(X_i) \right)$$

where *f* is a real measurable function defined on $\mathbb{R}$ such that

$$\mathbb{P}_G\left(|f|\right)=\int\left|f(x)\right|\mathrm{d}G(x)<\infty, \tag{8}$$

and

$$\mathbb{V}_G(f)=\int\left(f(x)-\mathbb{P}_G(f)\right)^2\mathrm{d}G(x)<\infty \tag{9}$$

It is known (see van der Vaart [7], pages 81-93) that $\mathbb{G}_n$ converges to a functional Gaussian process $\mathbb{G}$ with covariance function

$$\Gamma\left(\bar{f},\bar{\bar{f}}\right)=\int\left(\bar{f}-\mathbb{E}_G\left(\bar{f}\right)\right)\left(\bar{\bar{f}}-\mathbb{E}_G\left(\bar{\bar{f}}\right)\right)\mathrm{d}G(x) \tag{10}$$

at least in finite distributions. $\mathbb{G}_n$ is linear, that is, for f and g satisfying (9) and for $(a,b)\in\mathbb{R}^2$, we have

$$a\mathbb{G}_n(f)+b\mathbb{G}_n(g)=\mathbb{G}_n(af+bg).$$

This linearity will be useful for our proofs. We are now in position to state our main results.

2.2. Statements of Results

First introduce this notation for $\ell\geq 0$, $k\geq 2$, and $2\leq p\leq k$. Let f_i and g_i, $i=1,\cdots,k$ be C^1-functions with values in $\mathbb{R}$. Put $\mu_2=\sigma^2$ and $m_1=m$ and $h_\ell(x)=x^\ell$, $x\in\mathbb{R}$,

$$A(\ell)=h_\ell+\sum_{p=0}^{\ell-1}C_\ell^p(-1)^{\ell-p}\left(m_1^{\ell-p}h_p+(\ell-p)m_1^{\ell-p-1}m_ph_1\right) \tag{11}$$

$$B(p)=\sigma^{-(2p-1)}\left(A(2p-1)-\frac{1}{2}(2p-1)\sigma^{-2}\mu_{2p-1}A(2)\right) \tag{12}$$

$$C(p)=\sigma^{-2p}\left(A(2p)-p\sigma^{-2}\mu_{2p}A(2)\right) \tag{13}$$

and

$$D_k=\sum_{p=2}^k\left(f_p'(b_p)B(p)+g_p'(a_p)C(p)\right) \tag{14}$$

Here are our main results.

Theorem 1 *Let* $\mathbb{E}|X|^{4k}<\infty$, *for* $k\geq 2$. *Then*

$$T_n^*(f,g,k)=\sqrt{n}\left(T_n(f,g,k)-T(f,g,k)\right)\to\mathcal{N}\left(0,\sigma_k^2\right),$$

where

$$\sigma_k^2=\left(\int D_k^2(x)\mathrm{d}G(x)\right)-\left(\int D_k(x)\mathrm{d}G(x)\right)^2.$$

Corollary 1 (*Normality test*). *Let X be a* $\mathcal{N}\left(m,\sigma^2\right)$ *r.v. and let, for all* $k\geq 2$

$$T_k=\sum_{p=2}^k\left(f_p(0)+g_p\left(\frac{(2p)!}{2^p\,p!}\right)\right),$$

Then

$$\sqrt{n}\left(T_n(f,g,k)-T_k\right)\to\mathcal{N}\left(0,\sigma_{k,0}^2\right),$$

where

$$\sigma_{k,0}^2=\left(\int D_{k,0}^2(x)\mathrm{d}G(x)\right)-\left(\int D_{k,0}(x)\mathrm{d}G(x)\right)^2,$$

and

$$D_{k,0}=\sum_{p=2}^k\left(f_p'(0)B(p)+g_p'\left((2p)!/2^p\,p!\right)C(p)\right).$$

2.3. Particular Cases and Consequences

2.3.1. A General Test

Let G be an arbitrary *df* with a $4k^{\text{th}}$ finite moment for $k \geq 2$, this is $\int x^{4k} \mathrm{d}G(x) < +\infty$. We want to check whether a sample $X_1, \cdots, X_n$ is from G. We then select C^1-functions f_i and g_i, $i = 1, \cdots, k$ and compute the observed value $t_n^*(f, g, k)$ of $\sqrt{n}\left(T_n^*(f, g, k) - T^*(f, g, k)\right)$ and report the *p*-value of the test, that is $p = \mathbb{P}\left(\left|\mathcal{N}(0,1)\right| \geq \left|t_n^*(f, g, k)\right| s\right)$ where s^2 is either the exact variance σ_k^2 or its plug-in estimator

$$\hat{\sigma}_{k,n}^2 = \left(\frac{1}{n}\sum_{i=1}^n D_k^2\left(X_{j,n}\right)\right) - \left(\frac{1}{n}\sum_{i=1}^n D_k\left(X_{j,n}\right)\right)^2$$

Our guess is that using a greater value of k makes the test more powerful since the equality in distribution of univariate *r.v.*'s means equality of all moments when they exist (see page 213 in [8]). For $k = 2$, this result depends on the first eight moments. Then to find another *df* G_1 for which the p-value exceeds 5% would suggest it has the same eight moments as G, which is highly improbable. Simulation studies in Section 0 support our findings. Remark that we have as many choices as possible for the functions the $f_i's$ and $g_i's$.

Unfortunately, in the simulation studies reported below, we noticed that the plug-in estimator $\hat{\sigma}_{k,n}^2$ may hugely over estimate the exact variance and leads to accepting any data to follow that model, or significantly underestimate it and leads to reject data from the model itself. This is why we only use the exact variance here.

Now let us show how to derive chi-square tests from Theorem 1.

2.3.2. Generalized JB Test and Tests for Symmetrical *df*'s

Suppose that X is a symmetrical distribution. We have from Theorem 1 that

$$\sqrt{n}\left(\left(b_{n,p} - b_p\right), \left(a_{n,2} - a_p\right)\right) = \left(\mathbb{G}_n\left(B(p)\right), \mathbb{G}_n\left(C(p)\right)\right) + o_{\mathbb{P}}(1) \tag{15}$$

Since X is symmetrical, that is $\mu_{2\ell-1} = 0$ for $\ell \geq 1$, we may without loss of generality suppose that $m_1 = 0$ since replacing X by $X - m_1$ does affect neither the $\left(b_{n,p}, a_{n,p}\right)'s$ nor the $\left(b_p, a_p\right)'s$. Then we have from (11) and (12) that

$$C(p) = \sigma^{-(2p-1)} A(2p-1) = \sigma^{-(2p)}\left(h_{2p} - p\sigma^{-2}\mu_{2p}h_2\right)$$

and

$$B(p) = \sigma^{-(2p-1)}\left(h_{2p-1} - (2p-1)m_{2(p-1)}h_1\right)$$

By reminding that $h_p h_q = h_{p+q}$ for $p \geq 0$ and $q \geq 0$, we observe that the product $B(p) \times C(p)$ only includes functions h_j with odd $j's$ and then $\mathbb{E}\mathbb{G}_n\left(B(p) \times C(p)\right) = 0$. Thus

$$\sqrt{n}\left(\left(b_{n,p} - b_p\right), \left(a_{n,p} - a_p\right)\right) \to_d \mathbb{N}_2\left(0, \Sigma_p\right)$$

where $\left(\Sigma_p\right)_{11} = \mathbb{V}\mathrm{ar}\left(B(p)\right) = bj(p)$, $\left(\Sigma_p\right)_{22} = \mathbb{V}\mathrm{ar}\left(C(p)\right) = aj(p)$ and $\left(\Sigma_p\right)_{12} = 0$. We get

Corollary 2 *Let* $\int x^{4p}\mathrm{d}G(x) < \infty$ *for* $p \geq 2$ *and G be a symmetrical df. We have*

$$n\left(b_{n,p}^2 / bj(p) + \left(a_{n,p} - a_p\right)^2 / aj(b)\right) \to \chi_2^2. \tag{16}$$

For a standard normal random variable, we get $bj(2) = 6$ and $aj(2) = 24$ and the normality JB's test becomes a particular case of (16), which is a general chi-square test for an arbitrary *df* with $2p$-finite moments.

Corollary 3 *Let G be a Gaussian df. Then*

$$\frac{n}{6}\left(b_{n,2}^2 + \left(a_{n,2} - 3\right)^2 / 4\right) \to \chi_2^2.$$

We see that we obtain an infinite number of tests for the normality. For example, for $p = 3$, we have,

$$\frac{n}{360}\left(b_{n,3}^2 / 2 + \left(a_{n,3} - 15\right)^2 / 17\right) \to \chi_2^2, \text{ etc.}$$

2.4. A General Chi-Square Test

Consider (15) and put $abj(p) = \text{cov}(C(p), B(p))$ and suppose that $\Delta(p) = aj(p) \times bj(p) - abj(p)^2 \neq 0$. We have

Corollary 4 Let $\int x^{4k} dG(x) < \infty$ and $\Delta(p) \neq 0$ for $2 \leq p \leq k$. Then

$$\frac{n}{\Delta(p)} \left(aj(p) \left(b_{n,p} - b_p \right)^2 + bj(p) \left(a_{n,p} - a_p \right)^2 - 2 \times abj(p) \left(b_{n,p} - b_p \right) \left(a_{n,p} - a_p \right) \right)$$

converges in law to a χ_2^2 r.v..

It is now time to prove Theorem 1 before considering the simulation studies.

2.5. Proofs

Since *G* has at least first 4*k* moments finite, we are entitled to use the finite-distribution convergence of the empirical function process $\mathbb{G}_n$ as below. Let us begin to give the asymptotic law of $\mu_{n,\ell}$. By denoting $h_\ell(x) = x^\ell$, we have

$$\mu_{n,\ell} = \sum_{p=0}^{\ell} C_\ell^p \left(-\bar{X} \right)^{\ell-p} \left(\frac{1}{n} \sum_{i=1}^{n} X_i^p \right)$$

$$= \sum_{p=0}^{\ell} C_\ell^p (-1)^{\ell-p} \left(m_1 + \frac{\mathbb{G}_n(h_1)}{\sqrt{n}} \right)^{\ell-p} \left(m_p + \frac{\mathbb{G}_n(h_p)}{\sqrt{n}} \right)$$

$$= \left(m_\ell + \frac{\mathbb{G}_n(h_\ell)}{\sqrt{n}} \right) + \sum_{p=0}^{\ell-1} C_\ell^p (-1)^{\ell-p} \left(m_1^{\ell-p} + (\ell-p) m_1^{\ell-p-1} \frac{\mathbb{G}_n(h_1)}{\sqrt{n}} + o_p \left(n^{-1/2} \right) \right) \times \left(m_p + \frac{\mathbb{G}_n(h_p)}{\sqrt{n}} \right)$$

$$= m_\ell + h_\ell + \sum_{p=0}^{\ell-1} C_\ell^p (-1)^{\ell-p} \left(m_1^{\ell-p} m_p + \frac{\mathbb{G}_n(A_\ell)}{\sqrt{n}} \right) + o_p \left(n^{-1/2} \right),$$

where $A(\ell)$ is defined in (11) and where we used that the linearity of the empirical functional process. By observing that $\mu_\ell = \sum_{p=0}^{\ell} C_\ell^p (-m_1)^{\ell-p} (m_p)$, we finally obtain

$$\sqrt{n} \left(\mu_{n,\ell} - \mu_\ell \right) = \mathbb{G}_n (A(l)) + o_p(1). \tag{17}$$

Now the law of $b_{n,p}$ is given by

$$\sqrt{n} \left(b_{n,p} - b_p \right) = \frac{1}{\mu_{n,2}^{(2p-1)/2}} \sqrt{n} \left(\mu_{n,2p-1} - \mu_{2p-1} \right) - \frac{\mu_{2p-1}}{\mu_{n,2}^{(2p-1)/2} \mu_2^{(2p-1)/2}} \sqrt{n} \left(\mu_{n,2}^{(2p-1)/2} - \mu_2^{(2p-1)/2} \right)$$

By the delta-method, we have

$$\mu_{n,2}^{(2p-1)/2} = \left(\mu_2 + \frac{\mathbb{G}_n(A(2))}{\sqrt{n}} \right)^{\frac{2p-1}{2}} + o_p \left(n^{-1/2} \right)$$

$$= \mu_2^{\frac{2p-1}{2}} + \frac{2p-1}{2} \mu_2^{\frac{2p-3}{2}} \frac{\mathbb{G}_n(A(2))}{\sqrt{n}} + o_p \left(n^{-1/2} \right),$$

and then

$$\sqrt{n} \left(\mu_{n,2}^{(2p-1)/2} - \mu_2^{(2p-1)/2} \right) = \left(\frac{2p-1}{2} \right) \mu_2^{\frac{2p-3}{2}} \mathbb{G}_n(A(2)) + o_p(1)$$

and next, by noticing from 17 that $\mu_{n,\ell} \to \mu_\ell$ for all $\ell \leq 2k$,

$$\sqrt{n} \left(b_{n,p} - b_p \right) = \mathbb{G}_n \left(\sigma^{-(2p-1)} A(2p-1) - \frac{1}{2} (2p-1) \sigma^{-(2p-1)} \mu_{2p-1} A(2) \right) + o_p(1),$$

$$\mathbb{G}_n(B(p)) + o_p(1) \to \mathbb{G}(B(p)),$$

where $B(p)$ is given in (12). By the very same methods, we have

$$\sqrt{n}\left(a_{n,p}-a_{p}\right)=\mathbb{G}_{n}\left(C(p)\right)+o_{p}(1),$$

$C(p)$ is stated in (13). The delta-method also yields

$$\sqrt{n}\left(T_{n}(f,g,k)-T(f,g,k)\right)=T_{n}^{*}(f,g,k)$$

$$=\sum_{p=2}^{k}\left(f_{p}\left(b_{n,p}\right)-f\left(b_{p}\right)\right)+\sum_{p=2}^{k}\left(g_{p}\left(a_{n,p}\right)-g\left(a_{p}\right)\right)$$

$$=\sum_{p=2}^{k}f_{p}'\left(b_{p}\right)\mathbb{G}_{n}\left(B(p)\right)+\sum_{p=2}^{k}g_{p}'\left(a_{p}\right)\mathbb{G}_{n}\left(C(p)\right)+o_{p}(1)$$

$$=\mathbb{G}_{n}\left(\sum_{p=2}^{k}\left(f_{p}'\left(b_{p}\right)B(p)+g_{p}'\left(a_{p}\right)C(p)\right)\right)+o_{p}(1)$$

$$=\mathbb{G}\left(D_{k}\right)+o_{p}(1).$$

This completes the proof of the theorem. The proof of the corollary is a simple consequence of the theorem.

3. Simulation and Applications

3.1. Scope the Study

We want to focus on illustrating how performs the general test for usual laws such as Normal and Double Gamma ones. It is clear that the generality of our results that are applicable to arbitrary *d.f.*'s with some finite k^{th}-moment $(k\geq 2)$ deserves extended simulation studies for different classes of *df*'s. We particularly have to pay attention to the choice of k and of the functions *fi* and *gi*, depending on the specific model we want to test.

In this paper, we want to set a general and workable method to simulate and test two symmetrical models. The normal and the double-exponential one with density $f(x)=(\lambda/2)\exp(-\lambda|x|)$. We expect to find a test that accepts normality for normal data and rejects double-exponential data and to confirm this by the Jarque-Berra test, and to have another test that exactly does the contrary.

Once these results are achieved, we would be in position to handle a larger scale simulation research following the outlined method. Specially, fitting financial data to the generalized hyperbolic model is one the most interesting applications of our results.

3.2. The Frame

We first choose all the functions f_i equal to f_0 and all the functions g_i equal to g_0. We fix $k=3$, that is we work with the first twelve moments. As a general method, we consider two *df*'s G_1 and G_2. We fix one of them say G_1 and compute $T(f,g,k)=T(f,g,k,G_1)$ and the variance σ_k^2 from the exact distribution function G_1. We generate samples of size n from one the *df*'s (either G_1 or G_2) and compute $T_n(f,g,k)$. We repeat this B times and report the mean value t^* of the replicated values of $T_n^*=\sqrt{n}\left(T_n(f,g,k)-T(f,g,k)\right)/\sigma$ and report the p-value $p=\mathbb{P}\left(\left|\mathcal{N}(0,1)\right|\geq t^*\right)$. The simulation outcomes will be considered as conclusive if p is high for samples from G_1 and low for samples from G_2. The results are compared with those given by the Kolmogorov-Smirnov test (KST) and when the data are Gaussian, they are compared with the outcomes from JB's classical test.

3.3. The Results

We consider the following cases: G_1 is a Gaussian r.v $\mathcal{N}\left(m,\sigma^2\right)$; G_2 is double-exponential law $\mathcal{E}_d(\lambda)$ with density probability $f_2(x)=(\lambda/2)\exp(-\lambda|x|)$ and G_3 is a double-gamma law $\gamma_d(a,b)$ with probability density $f_3(x)=b^a/(2\Gamma(a))|x|^{a-1}\exp(-b|x|)$.

Normal Model $N\left(m,\sigma^2\right)$

The choice $f_0(x)=g_0(x)=x^2$ is natural since the Jarque-Berra test may be derived for our result for these functions and for $k=2$. The model is determined by these following parameters:

$(b_p, a_p), 2 \le p \le 6$	$T(f,g,k)$	σ
$(0,3),(0,15),(0,105),(0,946),(0,10395)$	234	500.2918

We recall that the variance of our statistic depends on the first $4k$ moments.

Simulation study.

Testing the model with $\mathcal{N}(0,1)$ data gives the following outcomes for $n = 20$

	$T_n(f,g,k)$	T_n^*	$p\%$	JB	$pJB\%$	KS	$pKS\%$
$N(0,1)$	232.16	−0.023	49.05	1.338	51.5	0.7709	23.35

and for $n = 100$,

	$T_n(f,g,k)$	T_n^*	$p\%$	JB	$pJB\%$	KS	$pKS\%$
$N(0,1)$	249.21	0.42	33.82	1.73	42.22	0.918	15.60

and for $n = 1000$,

	$T_n(f,g,k)$	T_n^*	$p\%$	JB	$pJB\%$	KS	$pKS\%$
$N(0,1)$	243.34	0.59	27.73	2.08	35.38	0.98	12.62

where *JB* is the classical Jarque-Berra statistic, *pJB* is the p-value of the *JB* test, *KS* is the Kolmogorov-smirnov statistic and *pKS* is the related p-value. Our model accepts the normality and this is confirmed by JB's test and by the Klmogorov-Smirnov test (KST). The simulation results are very stable and constantly suggest acceptance.

Testing the double-exponential versus the normal model.

Recall that the values (b_p, a_p) for $2 \le p \le 6$ are $(0,3)$, $(0,15)$, $(0,946)$, $(0,10395)$. Comparing these values with those of a normal model, it is natural to think that the test will fail since only the b_p coincide and the test is only based on the moments. Indeed, using data from $\mathcal{E}_d(1)$ gives for $n = 11$

	$T_n(f,g,k)$	T_n^*	$p\%$	JB	$pJB\%$	KS	$pKS\%$
$\mathcal{E}(1)$	411.25	1.81	3.47	1.98	37.98	0.91	15.67

and for $n = 22$

	$T_n(f,g,k)$	T_n^*	$p\%$	JB	$pJB\%$	KS	$pKS\%$
$\mathcal{E}(1)$	1624	18.70	0	6.43	4.09	0.9	15

Our test rejects the $\mathcal{E}_d(1)$ model for $n = 11$ and JB's test rejects it only for $n \ge 22$. We see here the advantage brought by the value $k = 3$ in our statistic. The KST has problems in rejecting the false $\mathcal{E}_d(1)$ even for $n = 1000$ that of Jarque-Berra.

Testing the double-gamma versus the normal model.

Let use $\gamma_d(a,b)$ data with $a_0 = (1+\sqrt{13})/2$ and $b = 1$. We have the outcomes for $n = 11$

	$T_n(f,g,k)$	T_n^*	$p\%$	JB	$pJB\%$	KS	$pKS\%$
$\mathcal{E}(1)$	527.8	3.09	0.099	4.22	12.5	0.99	12.45

and for $n = 22$

	$T_n(f,g,k)$	T_n^*	$p\%$	JB	pJB%	KS	pKS%
$\mathcal{E}(1)$	1055	10.16	0	6.41	4.12	0.99	11

We have similar results. Ou test rejects the $\mathcal{E}_d(1)$ model for $n = 12$ and JB's test rejects it only for $n \geq 18$. We see here the advantage brought by the value $k = 3$ in our statistic. Although the first four moments of a $\gamma_d(a_0,1)$ are 0, 1, 0 and 3, that is, the same of those of standard normal *rv*, this model is rejected. We already pointed out that the coefficients 4 and 6 are in fact based on the first eight moments and the discrepancy of moments higher than 4 results in the rejection.

Analysing the tables above, we conclude that our test performs better the JB's test against a double-gamma *df* with same skewness and kurtosis than a normal *df* for small sample sizes around ten and this is real advantage for small data sizes. Even for $k = 2$, our test is performant for the small values $n = 11$ and $n = 12$.

Double-exponential model $\mathcal{E}_d(\lambda)$.

We point out that the statistic $T_n(f,g,k)$ does not depend on the λ. Then we only consider $\lambda = 1$ in the following. We always use $f_0(x) = g_0(x) = x^2$. The model is determined by the following values.

$(b_p,a_p),\ 2 \leq p \leq 6$	$T(f,g,k)$	σ
$(0,6),(0,90),(0,2520),(0,113400),(0,7484400)$	8136	73473

Here, we do not have the Jarque-Berra test to confirm the results.

Simulation. Testing the model with $\mathcal{E}_d(\lambda)$ data gives the following outcomes, for $n = 800$.

	$T_n(f,g,k)$	T_n^*	$p\%$
$\mathcal{E}_d(1)$	7858,0176	−0.41	41,730

The simulation results are very stable and constantly suggest acceptance.

Testing normal data. Using normal data gives

	$T_n(f,g,k)$	T_n^*	$p\%$
$\mathcal{N}(0,1)$	236.019	−3.044	0.11

The $\mathcal{N}(0,1)$ model is rejected. We noticed that the rejection of normal data is automatically obtained for large sizes here, when *n* is greater than 900. For *n* between 500 and 900, rejection is frequent but acceptance occurs now and then. We also noticed that the variance of T_n^* are high and do not allow to reject normal data for small sizes. This leads us to consider other functions. Now consider the classes of functions

$$\theta u + (1+u^p)^p,\ p\ \text{even}\ .$$

We obtain good results for $n = 150$ with $\theta = 0.1$ and $p = 2$. In this case, the exact value of the statistic is 11.600. The double-exponential $\mathcal{E}_d(1)$ model is confirmed according to the following table

	$T_n(f,g,k)$	T_n^*	$p\%$
$\mathcal{E}_d(1)$	7.968	−0.7973	21.38

while the normal model is rejected as illustrated below:

	$T_n(f,g,k)$	T_n^*	$p\%$
$\mathcal{N}(0,1)$	3.001	−1.87	3.01

It is important to mention here that the KST is very powerful is rejecting the normal model with double-exponential and double-gamma data with extremely low p-value's.

3.4. Conclusion and Perspectives

We propose a general test for an arbitrary model. The methods are based on functional empirical processes theory that readily provides asymptotic laws from which statistical tests are derived. They depend on an integer *k* such that the pertaining *df* has 4*k* first finite moments. We get two kinds of tests. A general one based on functions f_i and g_i, $i = 1, \cdots, k$, with an asymptotic normal law. We derive from these results chi-square tests that are valid for general *df*'s and that includes the Jarque-Berra test of normality. Both tests use arbitrary moments. We only undergo simulation studies for the first kind of test. Our simulation studies show high performance for normality against other symmetrical laws such as double-exponential or double-gamma ones. For suitable choices of f_i, g_i, and *k*, the test performs well for small samples ($n = 20$) both for accepting the normal model and rejecting other models. We also show that for suitable choice of f_i and g_i, the test for the double-exponential model is also successful, but for sizes greater that $n = 150$. In upcoming papers, we will focus on detailed results on specific models and try to found out, for each case, suitable value of the parameters of the tests ensuring good performances for small data. A paper is also to be devoted to simulation studies for the khi-square tests and their applications to financial data.

References

[1] Jarque, C.M. and Bera, A.K. (1987) A Test for Normality of Observations and Regression Residuals. *International Statistical Review*, **55**, 163-172. http://dx.doi.org/10.2307/1403192

[2] McNeil, A.J., Frey, R. and Embrechts, P. (2005) Quantitative Risk Management: Concepts, Techniques and Tools. Princeton University Press, Princeton and Oxford.

[3] Bowman, K.O. and Shenton, L.R. (1975) Omnibus Contours for Departures from Normality Based on $\sqrt{b1}$ and b2. *Biometrika*, **62**, 243-250.

[4] D'Agostino, R.B. (1971) An Omnibus Test for Normality for Moderate and Large Size Samples. *Biometrika*, **58**, 341-348. http://dx.doi.org/10.1093/biomet/58.2.341

[5] D'Agostino, R.B. and Pearson, E.S. (1973) Tests for Departure from Normality: Empirical Results for the Distributions of b2 and $\sqrt{b1}$. *Biometrika*, **60**, 613-622.

[6] D'Agostino, R.B. and Tietjen, G.L. (1973) Approaches to the Null Distribution of $\sqrt{b1}$. *Biometrika*, **60**, 169-173. http://dx.doi.org/10.1093/biomet/60.1.169

[7] van der Vaart, A.W. and Wellner, J.A.(1996) Weak Convergence and Empirical Processes with Applications to Statistics. Springer, New-York.

[8] Loeve, M. (1977) Probability Theory I. Springer-Verlag, Berlin.