HERMITE NORMALITY TESTS

David Declercq *Patrick Duvaut*

ENSEA-ETIS 6, avenue du PONCEAU 95014, Cergy-Pontoise, France
e-mail : declercq@ensea.fr

ABSTRACT

This paper introduces a new test statistic of Normality which evaluates the cross covariances between choosen Hermite polynomials which are zero under the null hypothesis. The special form of the test leads to a modified sphericity statistic and we have called it *Hermite Normality Test* (S_H). We present briefly its asymptotical distribution both under the null and nonnull hypothesis. Large simulations have been made to compare some specific Hermite tests to three other taken in the litterature. If our test is better for a lot nonnormal populations but works worse for some other, the main point is that we defined in fact a wide range of tests which may match different nonnormal distributions.

1. INTRODUCTION

Many tests of normality from various types have been largely studied in the past fifty years. Mardia [1] has written a nearly exhaustive paper which collects the most famous tests of univariate and multivariate normality.
The main purpose of our work is to build a new test of departure from normality using the orthogonality of Hermite polynomials of gaussian variables. For more information about Hermite polynomials, refer to the book of Erdelyi [2].
Our approach which is described in the second section is to consider first a vector of collected Hermite polynomials. Other works dealing with normality tests based on nonlinearities of a gaussian variable can be found in [3]. Then departure from normality means that the covariance matrix of this vector becomes different from the identity matrix, which might be written as a sphericity test. It is important to note that we define a class of tests rather than a single one since we may combine different numbers and types of Hermite polynomials in order to match the specific statistical properties of the underlying nonnormal population. Some theoretical asymptotical results are also elaborated. The third section is dedicaced to the simulation study. We have compared the Hermite normality test to three omnibus tests: the D'Agostino test, the Anderson-Darling test and an omnibus test based on the skewness and the kurtosis. We have followed the procedure adopted by Pearson, D'Agostino and Bowman [4] and apply the different tests to a wide range of nonnormal distributions. We finally discuss the results and precise the natural advantages and possible extentions of the Hermite normality test.

2. THE HERMITE NORMALITY TESTS

2.1. definition

Let x be a standart gaussian variable $\mathcal{N}(0,1)$; we then build a sample p-variate vector $\boldsymbol{X}$ by applying polynomial nonlinearities of type Hermite up to order p. The normalisation of that vector gives the following form

$$\boldsymbol{X} \triangleq \left[H_1(x), \frac{H_2(x)}{\sqrt{2}}, \dots, \frac{H_p(x)}{\sqrt{p!}} \right]^{\mathsf{T}} \tag{1}$$

Because Hermite polynomials of gaussian variables are orthogonals for the expectation, the vector defined by (1) is zero mean and spherical: its covariance matrix $\boldsymbol{\Sigma}$ is equal to the identity.

$$\begin{cases} \boldsymbol{\mu} = E[\boldsymbol{X}] = \boldsymbol{0} \\ \boldsymbol{\Sigma} = Var[\boldsymbol{X}] = \boldsymbol{I}_p \end{cases} \tag{2}$$

A sphericity test was introduced by Mauchly [5] in order to test if a multivariate normal random vector is spherical or not. Replacing the normal vector by (1) leads to a modified sphericity that tests the univariate normality of the variable x.
Therefore, the following property

$$\boldsymbol{\Sigma} = \boldsymbol{I}_p \iff x \ is \ \mathcal{N}(0,1) \tag{3}$$

leads to build the Hermite Normality Tests with the statistic which is defined in terms of the sample covariance matrix $\boldsymbol{R} = \hat{\boldsymbol{\Sigma}}$ of $\boldsymbol{X}$ by

$$S_H \triangleq \frac{|\boldsymbol{R}|}{\left(\frac{Tr(\boldsymbol{R})}{p}\right)^p} \tag{4}$$

where $|\boldsymbol{R}|$ is the determinant of $\boldsymbol{R}$ and $Tr(\boldsymbol{R})$ its trace.
The test statistic above defined takes its values in the range $0 \le S_H \le 1$ and is asymptotically equal to 1 under the null hypothesis. It has then to be compared with a threshold between 0 and 1 to decide whether the tested sample is normal or not.

$$S_H \gtrless^{\mathcal{H}_0}_{\mathcal{H}_1} \eta \qquad \eta \in [0,1] \tag{5}$$

Hence, this defines a class of tests of departure from normality since we can choose the size of the vector $\boldsymbol{X}$ and the Hermite polynomials contained in the vector; for instance, $\left[H_1(x), \frac{1}{\sqrt{6}}H_3(x)\right]^{\mathsf{T}}$ defines a specific test $S_H^{(1,3)}$.

2.2. asymptotical performances under $\mathcal{H}_0$

Many works have been achieved on the sphericity test when the underlying sample vector is $\mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, see [6] and in a recent paper [7] Mokkadem applies the sphericity to test the whiteness of regular time series which are not necessary normal.

Its asymptotical distribution is known because of some results on the Wishart distribution of the sample covariance matrix of a mutivariate normal process. But when the spherical vector is nonnormal - our case - its sample covariance is not distributed as Wishart and the distribution of S_H is not known.

We then have to make use of limit theorems to find the asymptotical distribution of the Hermite Normality test. Starting from Borovkov work [8], we have proved the following statements both under the null hypothesis and the alternative one :

theorem 1 *if* $x \in \mathcal{N}(0,1)$, *the Hermite Normality statistic has the limit distribution*

$$N(1 - S_H(x))) \in \frac{1}{2}\sum_{i=1}^{p}\xi_{ii}^2 - \frac{1}{2p}\left(\sum_{i=1}^{p}\xi_{ii}\right)^2 + \sum_{i \neq j}\xi_{ij}^2$$

where the ξ_{ij} *are normaly distributed with zero mean and covariance matrix depending on the Hermite polynomial degrees.*

theorem 2 *if* $x \in f(x)$ *is nonnormaly distributed, we have*

$$S_H \xrightarrow{a.s.} m_{\mathcal{H}_1}$$

and

$$\sqrt{N}(S_H(x) - m_{\mathcal{H}_1}) \in \xi$$

where ξ *is centered normal with variance* $\sigma_{\mathcal{H}_1}$.
The value of the indeterminates $m_{\mathcal{H}_1}$ *and* $\sigma_{\mathcal{H}_1}$ *are directly linked to cumulants of the nonnormal distribution and the degrees of the Hermite polynomials considered.*

The proofs of those theorems can be found in [9] with the expressions of the indeterminates. S_H is then asymptotically distributed as a quadratic form of centered normal variates with a $\frac{1}{N}$ rate of convergence under $\mathcal{H}_0$, when it is centered normal with a rate in $\frac{1}{\sqrt{N}}$ under $\mathcal{H}_1$.

3. SIMULATIONS

3.1. Comments on our approach

Extensive studies have been achieved for the comparison of power of various tests [1] [4] [10]. Therefore, we have choosen to apply the nonnormal populations only to three tests except the Hermite normality ones:
(i) one based on the empirical distribution function (EDF), namely the Anderson-Darling test A^2,
(ii) the D'Agostino D test which depends on order statistics [11]
(iii) and finally an omnibus test K^2 based on the third an fourth order cumulants proposed by D'Agostino and Pearson [12].
These statistics are completely described in [13].
We now apply the four omnibus tests for 22 nongaussian populations of length N at the $100\varepsilon\%$ level of significance, that is to say ε is the false alarm probability under the gaussian hypothesis.

$$\varepsilon = Prob\{S_H(x) < \eta \ / \ \mathcal{H}_0\}$$

All the alternative distributions are well known [14] ; $S\alpha S$ is a symetric α-stable law and means for generating such a variate are described in [15]. We have gathered the distributions that are far from the gaussian density in table 2 to apply the tests on small samples ($N = 20$), the ones which are close to the normal in table 4 with $N = 100$ because for smaller sample sizes, the powers were too meaningless, and the other were considered with moderate sample sizes ($N = 50$).
In order to evaluate the powers of the tests, we count - for a particular alternative - the number of samples for which the value of the test lay beyond the $100\varepsilon\%$ level quantiles from a normal distribution. For that purpose, table 1 gives the 5% level quantiles (η) of six special Hermite normality tests for $N = \{20, 50, 100\}$ estimated with 500000 trials. More complete tables will be availiable in future papers.
We have considered 5000 samples of each nongaussian population and the total count divided by 50 gives the estimated power of each test for the different alternatives.
The same procedure was carried out at levels $\varepsilon \in \{0.01, 0.05, 0.1\}$ each for $N = \{20, 50, 100\}$, but for obvious reasons of space, we present on tables 2-4 only the results at the 5% significance level.

3.2. Discussion of the results

First of all, we want to justify the choice of the alternative tests considered.
Our intention was to pick up one of the most powerfull omnibus test out of three different classes. The comments made in [1] and [4] lead us to do this choice. we could have replaced A^2 by the famous Shapiro-Wilkes W test, but Dyer [16] consider that they have nearly similar power.
For each nonnormal population within tables 2-4, we precise their normalised moments in order to quantify their distance from the normal distribution

$$\begin{cases} \sqrt{\beta_1} = \sqrt{\frac{\mu_3^2}{\mu_2^3}} & = 0 \ \ if \ \ x \in \mathcal{N}(0,1) \\ \beta_2 = \frac{\mu_4}{\mu_2^2} & = 3 \ \ if \ \ x \in \mathcal{N}(0,1) \end{cases}$$

where μ_i are the moments about the mean.
To ease the analysis of the results, we have emphasized the most powerfull test in bold.
Our aim is to not compare the alternative tests together, but with the six Hermite normality tests considered. We note although that the Anderson-Darling test is better for skewed alternatives when for symmetrical nonnormal alternatives, K^2 is more powerfull for platykurtic populations ($\beta_2 < 3$) and D for leptokurtic ones ($\beta_2 > 3$).
It can be noticed, in the ligth of the results in table 2 that :

- concerning the whole Hermite normality test for symmetrical populations, S_H performs better for platykurtic than for long tail ones since $S_H^{(1,3)}$ and $S_H^{(1,2,3)}$ are the most powerfull while $\beta_2 \leq 3$. On the contrary,

N	$S_H^{(1,2)}$	$S_H^{(1,3)}$	$S_H^{(1,8)}$	$S_H^{(2,3)}$	$S_H^{(2,4)}$	$S_H^{(1,2,3)}$
20	0.6180	0.2437	0.4355	0.4143	0.1253	0.1508
50	0.8017	0.4893	0.5617	0.4936	0.1975	0.3212
100	0.8872	0.6560	0.6389	0.6005	0.2985	0.4753

Table 1. 5% quantiles of the Hermite normality S_H test

population	$\sqrt{\beta_1}$	β_2	A^2	K^2	D	$S_H^{(1,2)}$	$S_H^{(1,3)}$	$S_H^{(1,8)}$	$S_H^{(2,3)}$	$S_H^{(2,4)}$	$S_H^{(1,2,3)}$
symmetrical											
beta(0.5, 0.5)	0	1.5	64	71	1	63	78	4	1	2	**87**
S_B ($\gamma = 0, \delta = 0.5$)	0	1.63	38	57	0	39	**69**	4	0	2	**69**
t ($\nu = 1$)	0	×	83	79	**84**	73	58	63	71	45	72
skewed											
χ^2 ($\nu = 2$)	2	9	78	52	52	**84**	12	34	39	10	60
lognormal	6.2	114	90	74	76	**93**	29	51	59	25	79
$S\alpha S$ ($\alpha = 0.5$)	×	×	**100**	98	**100**	94	90	84	90	72	96
Cauchy : $S\alpha S$ ($\alpha = 1$)	×	×	77	74	**78**	69	54	60	69	43	69

Table 2. Estimation of power based on 5000 samples of size $N = 20$ for $\varepsilon = 0.05$

population	$\sqrt{\beta_1}$	β_2	A^2	K^2	D	$S_H^{(1,2)}$	$S_H^{(1,3)}$	$S_H^{(1,8)}$	$S_H^{(2,3)}$	$S_H^{(2,4)}$	$S_H^{(1,2,3)}$
symmetrical											
uniform(*beta*(1, 1))	0	1.8	60	87	58	59	**95**	28	0	59	90
Tukey ($\lambda = 2.5$)	0	1.9	38	73	56	37	**89**	32	0	64	74
Laplace	0	6	55	45	**60**	37	26	25	33	14	36
t ($\nu = 4$)	0	14	44	48	**54**	43	33	31	40	24	43
t ($\nu = 2$)	0	×	89	85	**92**	79	74	66	67	58	78
skewed											
S_B ($\gamma = 1, \delta = 1$)	0.7	2.9	71	34	13	**80**	7	36	18	2	71
χ^2 ($\nu = 10$)	0.9	4.2	50	43	26	**65**	14	26	41	11	53
Erlang ($\nu = 2$)	1.42	6	90	75	58	**96**	28	43	64	21	92

Table 3. Estimation of power based on 5000 samples of size $N = 50$ for $\varepsilon = 0.05$

population	$\sqrt{\beta_1}$	β_2	A^2	K^2	D	$S_H^{(1,2)}$	$S_H^{(1,3)}$	$S_H^{(1,8)}$	$S_H^{(2,3)}$	$S_H^{(2,4)}$	$S_H^{(1,2,3)}$
symmetrical											
beta(2, 2)	0	2.1	33	72	72	31	**90**	50	0	86	59
S_U ($\gamma = 0, \delta = 3$)	0	3.5	11	15	**17**	14	10	5	15	7	15
logistic	0	4.2	25	31	**38**	27	22	10	25	13	28
skewed											
beta(2, 3)	0.3	2.4	41	46	31	48	44	**52**	1	29	51
Weibull ($k = 2$)	0.6	3.25	62	53	14	**80**	13	37	47	8	75
$S\alpha S$ ($\alpha = 1.7$)	×	×	76	83	**85**	81	79	65	74	72	81
$S\alpha S$ ($\alpha = 1.9$)	×	×	29	40	39	38	37	28	39	33	**41**

Table 4. Estimation of power based on 5000 samples of size $N = 100$ for $\varepsilon = 0.05$

the Hermite normality test seems to work rather bad for very long tail populations as it can be seen for the *logistic*, the *cauchy* or the *student's t* distributions.

- for the skewed distributions, S_H - and specially $S_H^{(1,2)}$ - is almost ever the most powerfull, but its domination decreases as much as the tails of the concerned population becomes heavier.
- we moreover remark that $S_H^{(2,3)}$ and $S_H^{(2,4)}$ never surpasses the other tests, and we deduce that the use of H_1 in the construction of the test is decisive ; which may be explained by the fact that the normalisation of the samples to be tested forces the first term of the matrix $\boldsymbol{R}$ to 1 and tends to make the statistic more robust.
- The best advantage of the Hermite normality test is that we make use not of a unique test, but of a complete class of tests. Therefore, this engage us to say that for a special alternative, there exists an optimal set of Hermite polynomials that fit the test to the data. We can see an illustration in the fact that $S_H^{(1,8)}$ does not work particulary well except for $beta\,(2,3)$; and if we consider the gathered powers of $S_H^{(1,2)}$, $S_H^{(1,3)}$ and $S_H^{(1,2,3)}$, the Hermite normality test performs best in 13 cases out of the 22 nonnormal population tested. We are although convinced that other combinations of Hermite polynomials would perform better than the tests presented here for each nongaussian distribution.

4. CONCLUSION

We have introduced a new test of normality using the orthogonal property of Hermite polynomials weighted by a gaussian density and present it in the form of a modified sphericity test. Besides the plenty litterature concerning this test, the determination of its asymptotical distribution when we consider a nonnormal multivariate vector present some difficulties. However, with the use of a limit theorem, we have proved that the Hermite statistic is distributed as a quadratic form under the null hypothesis and is normal for the alternative one. The Hermite normality test has also been compared with three other tests through a wide range of nonnormal populations. in addition to the fact that for a lot of distributions S_H seems to be very powerfull, the leading comment is that one have a complete class of tests at his disposal and can therefore choose the optimal test for his job. Our future work will be to find the parameters of the nonnormal distribution that impose the choice of a particular test, that means to choose the bet set of polynomials when some prior known upon the datas.

REFERENCES

[1] K.V. Mardia. Tests of univariate and multivariate normality. *Handbook of Statistics*, (1):279–320, 1980.

[2] A. Erdelyi and Al. *Higher transcendental Functions, Vol. II.* Bateman Manuscript Project, McGraw-Hill, New York, 1953.

[3] K. Choukri. *Un formalisme pour les tests statistiques de conformité de modèles.* PhD thesis, ENST, 1993.

[4] E.S. Pearson, R.B. D'Agostino, and K.O. Bowman. Tests for departure from normality: Comparison of powers. *Biometrika*, (64):231–246, 1977.

[5] J.W. Mauchly. Significance test for sphericity of a normal n-variate distribution. *Ann. Math. Stat.*, (11):204–209, 1940.

[6] T.W. Anderson. *An Introduction to Multivariate Statistical Analysis.* Wiley, New York, 1958.

[7] A. Mokkadem. Sur quelques nouveaux tests pour les séries chronologiques. *C.R. Acad. Sci. Paris*, (318-I):755–758, 1994.

[8] A. Borovkov. *Statistique mathématique.* Mir, Moscou, 1987.

[9] D. Declercq and P. Duvaut. Asymptotic distribution of the hermite normality test. In *submitted to the Wokshop HOS*, Alberta, 1997.

[10] M.A. Stephens. Edf statistics for goodness of fit and some comparisons. *J. Amer. Statis. Assoc.*, (69):730–737, 1974.

[11] R.B. D'Agostino. An omnibus test of normality for moderate and large sample sizes. *Biometrika*, (58):341–348, 1971.

[12] R.B. D'Agostino and E.S. Pearson. Tests for departures from normality. empirical results for the distributions of b_2 and $\sqrt{b_1}$. *Biometrika*, (60):613–622, 1973.

[13] R.B. D'Agostino and M.A. Stephens. *"Goodness-of-Fit" Techniques.* Marcel Dekker INC vol. 68, 1986.

[14] A. Stuart Sir M. Kendall and J.K. Ord. *Kendall's Advanced Theory of Statistics - Vol. 1 (5th edition).* Griffin, London, 1987.

[15] C.L. Nikias and M. Shao. *Signal Processing with Alpha-stable Distributions and Applications.* John Wiley, Univ. South. CA, 1995.

[16] A.R. Dyer. Comparison of tests for normality with a cautionary note. *Biometrika*, (61):185–189, 1974.