Abstract

We present in this paper an extended overview of the Hermite Normality test. This test makes use of the Hermite polynomials and a modified sphericity statistic to determine whether a unidimensional, standardised and white sample is normal or not. Its major advantage is to yield not a single test but a real class of test statistics which allows us to match the normality test to the data. We give the limit distribution of the Hermite tests both for the null and nonnull hypothesis and especially for those built with two polynomials. We have determined the tests asymptotically the most powerful for some fixed alternative distributions and made extensive simulations to compare the Hermite tests with three other. The results are good and encourage us to go further with the generalisation of the Hermite test to correlated and multivariate data.

Nous présentons dans cet article une étude complète du test de normalité d’Hermite. Ce test utilise les propriétés des polynômes d’Hermite et une statistique de sphéricité modifiée pour décider si un échantillon monodimensionnel, standardisé et blanc est gaussien ou non. L’avantage majeur de cette approche est de définir une famille de statistiques qui va permettre d’adapter le choix d’un test particulier aux données. Nous avons établi la distribution asymptotique du test d’Hermite sous l’hypothèse nulle et sous l’hypothèse alternative et étudié en détails le cas particulier de tests à deux polynômes. Nous avons déterminé les tests asymptotiquement les plus puissants pour quelques distributions alternatives et effectué un grand nombre de simulations afin de comparer le test d’Hermite à trois autres tests. Les bons résultats obtenus nous encouragent à généraliser le test d’Hermite aux cas de données colorées et multivariées.

1 Introduction

This paper deals with testing normality of unidimensional and uncorrelated samples. Although a lot of work have been achieved on this topic over fifty years (see the paper of Mardia to have a complete overview [1]), there is a major problem for the statistician when no prior knowledge on the data are available: the choice of a test among the huge quantity proposed in the literature. Some studies like Hogg’s family of tests based on the kurtosis [2] and comparison of powers such as in [3] or [4] can give some advice. Nevertheless, if some tests work best for certain alternative hypothesis, when nothing is known about the samples, one can not claim that a test is more powerful than another and the problem of the choice of a test appears. The Hermite Normality test $S_H$ introduced in [5] try to answer to this question because it defines not a single test, but a complete class of tests, each provides different statistical behaviour. This modular property makes the choice of a test of normality easier because with the same mathematical structure, the Hermite test can match a lot of nonnormal populations. We recall in section 2 the means of generation of the Hermite normality statistic, in section 3 we make a complete study of the asymptotic distribution of our test introduced in [6]. The fourth section is devoted to the particular tests with only two polynomials which are very useful in experiments and the question of choosing a best Hermite test for a fixed alternative is studied. A normality test usually works with very small number of samples and the asymptotic behaviour may be inadequate. In the section 5 we fit the distribution of the Hermite test for finite
sample size $N$ with Pearson curves and mixture models. This paper ends up in section 6 with the calculation of the critical values of the Hermite test and a comparison study.

Remark

In sections 2-4, we have considered standardised random variates, which means that the asymptotical study of the Hermite statistic holds for samples with known mean and known variance. In order to deal with more general samples and to build a scale and location invariant test, we will consider for the nonasymptotic study and the comparison of power a pre-standardisation of the samples. We discuss the effects of such a transformation on the asymptotical behaviour in section 5.

2 The Hermite Normality statistic

The main idea that led us to build this new normality statistic is to test whether the cross-covariances between Hermite polynomials of standard Gaussian variables are zero or not. Actually, the property of statistical orthogonality of Hermite polynomials weighted by a normal density ensures that

$$E[H_i(x)H_j(x)] = i! \delta(i-j) \quad \iff \quad x \sim \mathcal{N}(0,1) \quad (1)$$

The first step is to build a vector which collects some normalised Hermite polynomials of a random variate $x$ with arbitrary distribution

$$X = [\hat{H}_{i_1}(x), \hat{H}_{i_2}(x), \ldots, \hat{H}_{i_p}(x)]^T \quad (2)$$

where $\hat{H}_i(x) = \frac{H_i(x)}{\sqrt{i!}}$ and all the subscripts $i_k$ are distinct. Using (1), when $x$ is normally distributed, the vector $X$ becomes spherical:

$$\left\{ \begin{array}{l}
\mu = E[X] = 0_p \\
\Sigma = Var[X] = I_p
\end{array} \right.
\quad (3)$$

The particular structure of this vector is helpful to derive a complete class of normality tests, linked to the choice of a number of polynomials $p$ and their degrees $i_k$. Testing the normality of the variate $x$ and the sphericity of the vector $X$ are also two equivalent problems. We then make use of the sphericity statistic introduced by Mauchly [7] and apply it to the vector defined by (2) to yield a new Normality test we have noted $S_H$. The two hypothesis we have to test are

$$\mathcal{H}_0 : \Sigma = I_p \quad \iff \quad x \sim \mathcal{N}(0,1)$$

$$\mathcal{H}_1 : \Sigma \neq I_p \quad \iff \quad x \not\sim \mathcal{N}(0,1)$$

and the Hermite normality statistic is

$$S_H = \frac{|R|}{(Tr(R))^{\frac{p}{2}}} \quad (4)$$

where $R = \hat{\Sigma}$ is the sample covariance matrix of $X$, $|R|$ its determinant and $Tr(R)$ its trace. The decision rule is then

$$S_H \sim \mathcal{H}_0 \iff \eta \quad \eta \in [0,1] \quad (5)$$

The threshold $\eta$ is determined by the distribution of $S_H$ under $\mathcal{H}_0$ and therefore depends on a fixed false alarm probability (denoted generally percentage point or quantile), the sample size and obviously the parameters of the vector (2).

This approach defines a complete class of normality tests while one can choose the number of Hermite polynomials and their degrees to build the vector (2). For example, $X = [\hat{H}_1(x), \hat{H}_3(x)]^T$ yield to a specific test we will note $S_H^{(1,3)}$. This emphasizes the fact that because we have many tests at our disposal, each based on the same mathematical structure but with different statistical behaviours, we can try to adapt the choice of the parameters to the alternative hypothesis when some prior knowledge is available.
3 Asymptotical distribution

3.1 Distribution under $\mathcal{H}_0$

This section is devoted to the asymptotic distribution of the Hermite Normality test, that is when the size $N$ of the samples tends to infinity. In this case, we must make use of limit theorems such as the central limit theorem. Although our test has a sphericity structure, we can not apply the results pointed out by Anderson [8] because they consider the sphericity statistic with a normal vector - and the one we consider (2) is not. We therefore have to introduce a limit theorem with tensorial notations derived with the help of two from Borovkov (9, pp. 44). One does not need to know everything on tensor algebra to understand the theorem and the sequel of our study, but we prefer recalling some important notations and relations:

(i) if $A$ is a matrix in $\mathbb{R}^p$, $A^*$ will denote its dual form, that is a morphism of $\mathcal{L}(\mathbb{R}^p; \mathbb{R})$ and $\otimes$ is the classical tensor product.

(ii) $\text{res} = A^* \otimes B$ is called contraction and its result is a scalar (a real).

$$\text{res} = \sum_{i=1}^{p} \sum_{j=1}^{p} a_{ij} b_{ij}$$

(iii) $\text{RES} = A \otimes B^*$ is a direct product and gives a $4^{th}$ order tensor, that is a four dimensionnal array:

$$\text{RES} = [\text{res}_{ij}^{kl}]_{1 \leq i, j, k, l \leq p}$$

(iv) the right product of $4^{th}$ order tensor by a matrix provides a matrix and the left product by a dual matrix gives a dual matrix.

$$\text{RES} \otimes C = D$$

$$C^* \otimes \text{RES} = D^*$$

Remark

We do not need in particular to rewrite the arrays of arbitrary order with the so-called Einstein conventions (notations with indices) which are highly useful for tensorial approaches. In our case, the gain is too small compared with the lack of readability. One can find a complete overview of the tensorial properties and definitions in [10].

Theorem 1 let $x$ be a random variate with the distribution $F_0$ and the two functions

$$G : \mathbb{R} \rightarrow \mathcal{M}(p, p) \quad h : \mathcal{M}(p, p) \rightarrow \mathbb{R}$$

where $\mathcal{M}(p, p)$ is the space of square real valued matrix of size $p$. Let furthermore $h(T)$ be continuous in $A = \int G(x)dF_0(x)$ and the tensor of covariance $\Sigma = \int (G - A) \otimes (G - A)^* dF_0(x)$ be finite. We have then the following result

$$S_N(x) = h\left(\frac{1}{N} \sum_{k=1}^{N} G(x_k)\right) \xrightarrow{a.s.} h(A)$$

$$\sqrt{N}(S_N(x) - h(A)) \in h'(A)^* \otimes \xi = \sum_{i,j=1}^{p} \sum_{j=1}^{p} \frac{\partial h(A)}{\partial t_{ij}} \xi_{ij}$$

where $h'(T)$ is the matrix of derivatives of $h(T)$ and $\xi \in \mathcal{N}(0, \Sigma)$ is a matrix whose components are centered normal with the tensor of covariance $\Sigma$. If $h'(A)^* \otimes \xi \xRightarrow{a.s.} 0$ and the tensor of the second order derivatives $h''(T)$ is finite in $A$, then

$$N(S_N(x) - h(A)) \leq 2 \xi^* \otimes h''(A) \otimes \xi$$

$$\leq \frac{1}{2} \sum_{i,j,r,s=1}^{p} \frac{\partial^2 h(A)}{\partial t_{ij} \partial t_{rs}} \xi_{ij} \xi_{rs}$$ (6)
The proof is straightforward while it is just an extension of Borovkov’s theorems. We have then to introduce some notations in order to apply this theorem to the Hermite test:

\[
\begin{align*}
&g_{ij}(x) = \tilde{H}_i(x)\tilde{H}_j(x) \quad \forall (i,j) \in \{1, \ldots, p\} \\
&G = [g_{ij}(x)] \text{ symmetrical matrix} \\
&A = E[G(x)] = [a_{ij}] \\
&R = \left[\frac{1}{n} \sum_k g_{ij}(x_k)\right] = \tilde{A} \\
&\Sigma = E[(G-A) \odot (G-A)^*] = [\sigma_{ij}^r] \\
&\sigma_{ij}^r = E[(g_{ij}(x) - E[g_{ij}(x)])(g_{rs}(x) - E[g_{rs}(x)])] \\
&\text{with } 1 < i,j,r,s < p \\
&S_H = h(R) = \frac{|R|}{(Tr(R))^{p/2}}
\end{align*}
\]

with these notations, we have proved the following theorem

**Theorem 2** Distribution of $S_H$ under the null hypothesis

*let $x \in N(0,1)$, the Hermite Normality statistic (4) verifies

$$S_H(x) \xrightarrow{a.s.} 1$$

and has the asymptotic behaviour

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

(8)

where $\xi = [\xi_{ij}] \in N(0, \Sigma)$

and the distinct terms of $\Sigma$ are

$$\{\sigma_{ii}^*, \sigma_{ij}, \sigma_{ij}^*, \sigma_{ii}^r, \sigma_{ij}^r\}$$

*proof*

because $x$ is a normal variable $N(0,1)$, we have $E[\tilde{H}_i(x)\tilde{H}_i(x)] = 1$ and $E[\tilde{H}_i(x)\tilde{H}_j(x)] = 0$ and then $A = I$.

We immediately deduce that

$$h(A) = \frac{|I|}{(Tr(I))^{p/2}} = 1 \quad \Rightarrow \quad S_H(x) \xrightarrow{a.s.} 1$$

Using the classical forms of the derivatives of the determinant and the trace of $G$ (we recall that $G$ is symmetrical)

$$\frac{\partial |G|}{\partial g_{ii}} = G_{ii} \quad \frac{\partial |G|}{\partial g_{ij}} = 2G_{ij} \quad \frac{\partial Tr(G)}{\partial g_{ii}} = 1 \quad \frac{\partial Tr(G)}{\partial g_{ij}} = 0$$

we calculate the matrix of derivatives $h'(G)$ and with the values of the cofactors $A_{ii} = 1$ et $A_{ij} = 0$ we found that

$$\frac{\partial h(A)}{\partial g_{ii}} = \frac{\partial h(A)}{\partial g_{ij}} = 0 \quad \Rightarrow \quad h'(A) = 0$$

We therefore cannot apply the first part of theorem 1 and the second part requires the expression of the tensor of second order derivatives $h''(G)$ which contains 7 types of terms. With $G_{ij}^{rs}$ the cofactor of $g_{rs}$ in $G_{ij}$, when applied to $A$, we have

$$A_{ii} = 1 \quad A_{ij} = 0 \quad A_{ij}^{rs} = 0 \quad A_{ij}^{ri} = -1$$
and the terms of $\mathbf{H}'(\mathbf{A})$ are

$$
\frac{\partial^2 h(\mathbf{A})}{\partial^2 g_{ii}} = -\frac{p+1}{p} \quad \frac{\partial^2 h(\mathbf{A})}{\partial g_{ii} \partial g_{jj}} = \frac{1}{p} \quad \frac{\partial^2 h(\mathbf{A})}{\partial^2 g_{ij}} = -2
$$

$$
\frac{\partial^3 h(\mathbf{A})}{\partial g_{ii} \partial g_{ij}} = \frac{\partial^3 h(\mathbf{A})}{\partial g_{ii} \partial g_{rs}} = \frac{\partial^3 h(\mathbf{A})}{\partial g_{ij} \partial g_{sr}} = \frac{\partial^3 h(\mathbf{A})}{\partial g_{ij} \partial g_{rs}} = 0
$$

consequently, the limit distribution of $S_H$ has the following expression

$$
\frac{1}{2} \left( \mathbf{\xi}^* \otimes \mathbf{H}'(\mathbf{A}) \otimes \mathbf{\xi} \right)
$$

$$
= \frac{1}{2} \left( \sum_{i=1}^{p} \frac{1-p}{p} \xi_{ii}^2 + \sum_{i \neq j}^{p} \frac{1}{p} \xi_{ii} \xi_{jj} - 2 \sum_{j > i} \xi_{ij}^2 \right)
$$

$$
= -\frac{1}{2} \sum_{i=1}^{p} \xi_{ii}^2 + \frac{1}{2p} \left( \sum_{i=1}^{p} \xi_{ii} \right)^2 - \sum_{j > i} \xi_{ij}^2
$$

Because the values of $S_H(x)$ are always within $[0, 1]$, we give the distribution of $N(1 - S_H(x))$ - instead of $N(S_H(x) - 1)$ in theorem 1 - and a sign shifting in the above equation ends up the proof.

We finally have to give the expressions of the terms in $\mathbf{\Sigma}$ obtained with the help of equations (22)-(26) that give the expectations of products of Hermite polynomials of normal variates. Because the tensor of covariance $\mathbf{\Sigma}$ is symmetrical, we consider only the terms regarding $i < j < r < s$.

$$
\sigma_{ii}^{ij} = \left( \sum_{k=0}^{i} \binom{i}{k} \binom{2k}{k} \right) - 1
$$

$$
\sigma_{ij}^{jj} = \left( \sum_{k=0}^{i} \binom{i}{k} \binom{j}{k} \binom{2k}{k} \right) - 1
$$

$$
\sigma_{ij}^{ij} = \sum_{k=0}^{i} \binom{i}{k} \binom{j}{k} \binom{2k}{k}
$$

$$
\sigma_{ii}^{rs} = \sum_{k=0}^{i} \frac{k!}{k!} \frac{\sqrt{s!}}{\sqrt{r!}} \binom{i}{k} \binom{r}{k} \binom{s}{k}
$$

with $k' = k - \left( \frac{s - r}{2} \right)$

$$
\sigma_{ij}^{rs} = \frac{1}{\sqrt{i!j!r!s!}} \min(i, i+j, i+s, i+j+r) \sum_{k=0}^{\min(i, i+j, i+s, i+j+r)} k!k'!(i+j-2k) \binom{i}{k} \binom{j}{k} \binom{2k}{k} \binom{s}{k'}
$$

with $k' = k + \frac{r + s - i - j}{2}$ and $s \leq i + j + r$

$\sigma_{ii}^{rs}$ and $\sigma_{ij}^{rs}$ are zero if respectively $(s - r)$ and $(r + s - i - j)$ are odd.

The Hermite Normality test in then asymptotically distributed as a quadratic form of centered and correlated normal variates. We know that for this type of distribution, there exists no close form for the probability density function, but one can express its characteristic function $[11]$. Let $Q$ be the canonical form of (8) whose rank is $q \leq \frac{(p+1)}{2}$ and $\{\xi'_1 \ldots \xi'_q\}$ its normal variates deduced from the $\xi_{ij}$ by linear transformations. $M$ will denote the coefficient’s matrix of $Q$ and $\Sigma'$ the covariance matrix of $[\xi'_1 \ldots \xi'_q]$. The characteristic function of $S_H$ under the null hypothesis is then

$$
\phi_{S_H}(z) = \prod_{k=1}^{q} \frac{1}{\sqrt{1 - 2iz\lambda_k}}
$$

where $\lambda_k$ are the nonnegative eigenvalues of $\Sigma'M$. We can therefore give the asymptotic cumulants of $S_H$:

$$
\kappa_n = 2^{n-1} (n-1)! \sum_{k=1}^{q} \lambda_k^n
$$
which means for the mean and the variance

\[ \mu_{H_0} = \sum_{k=1}^{q} \lambda_k, \quad \sigma_{H_0}^2 = 2 \sum_{k=1}^{q} \lambda_k^2 \]

Some numerical methods exist to approximate the p.d.f. \( f(x) \) corresponding to \( \phi_{S_{H_0}}(z) \) (cf. [11]), and we chose a simple approximation proposed by Rice ([12], p. 99): a Pearson type III curve (gamma distribution) with the same mean and variance

\[ g(x) = f(x) = \left( \frac{\mu_{H_0}}{\sigma_{H_0}^2} \right)^r \frac{x^{r-1}}{\Gamma(r)} e^{-\frac{x^2}{2\sigma_{H_0}^2}} \quad \text{with} \quad r = \frac{\mu_{H_0}^2}{\sigma_{H_0}^2} \]

We finally stress the fact that the rate of convergence of the Hermite normality test is \( \frac{1}{\sqrt{N}} \), when lots of normality statistics converge in \( \frac{1}{\sqrt{N}} \). Testing hypothesis is a detection problem and is usually based on the contrast between the behaviour of a statistic under both hypothesis. The rate of convergence of a test may be interpreted as the decrease’s speed of the variance around the expected values under the two hypothesis as the size of sample grows. In our case, the distribution of the Hermite statistic converges in \( \frac{1}{\sqrt{N}} \) under \( H_0 \) and \( \frac{1}{\sqrt{N}} \) under \( H_1 \) (see next section) - thereby, the contrast between the two hypothesis grows more rapidly than in the case of both convergence in \( \frac{1}{\sqrt{N}} \).

Besides that fact, the power of the test is not really affected by this rate of convergence because when a quantile is fixed at each sample size \( N \), the power of a test depends on its behaviour under the alternative hypothesis.

### 3.2 Distribution under \( H_1 \)

While taking a look at theorem 1, it is clear that the asymptotic behaviour of the Hermite Normality test under the nonnull hypothesis - that is \( x \) has a fixed and nonnormal distribution - is conditioned by the knowledge of \( A \) and \( \Sigma \). Those quantities can be obtained using relations (27)-(28). Actually, if we consider the Gram-Charlier expansion of the nonnormal p.d.f. \( f(x) \) (see [13], p. 222)

\[ f(x) = f_0(x) \sum_{j=0}^{\infty} c_j H_j(x) \]

where \( f_0(x) = N(0,1) \), we have

\[ \mathbb{E}[H_i(x)] = \int_{\mathbb{R}} H_i(x) f(x) dx = \sum_{j=0}^{\infty} c_j \int_{\mathbb{R}} H_i(x) H_j(x) f_0(x) dx = i! c_i \]

It is well known that \( c_i \) can be expressed with the moments or the cumulants of the distribution \( f(x) \) up to the \( i \)th-order. The ability to calculate \( \mathbb{E}[H_i(x)] \) and thereafter \( A \) and \( \Sigma \) which drive the behaviour of \( S_H \) under a nonnormal hypothesis depends effectively on the knowledge of its cumulants (moments) up to a sufficient order fixed by the Hermite polynomials choosen to build the test.

With the same notations as indicated by (7) for the theorem 2, we remark that the matrix of first order derivatives of \( h \) is no more zero in \( A \) (\( h'(A) \neq 0 \)) and the theorem 1 ensures that:

**Theorem 3** Distribution of \( S_H \) under the nonnull hypothesis

Let \( x \in f(x) \) a nonnormal random variate:

\[ S_H \xrightarrow{\alpha \xi} h(A) \]

and the limit distribution is

\[ \sqrt{N}(S_H(x) - h(A) ) \in h'(A) \otimes \xi \]

\[ \xi \in \mathcal{N}(0, \Sigma) \]

\[ S_H \text{ is asymptotically centered normal with variance } \sigma_{S_H}^2 = h'(A) \otimes \Sigma \otimes h'(A), \text{ the values of } A \text{ and } \Sigma \text{ are obtained with (27)-(28)}. \]
4 Examples for two polynomials

In order to have a better overview of the asymptotic properties of the Hermite test, we will give the explicit distributions for a very simple type of Hermite tests, highly used in simulations, when we consider only two polynomials, both under the null and a nonnull particular hypothesis: the uniform distribution.

4.1 Distribution under the normal hypothesis

Let \( x \in \mathcal{N}(0, 1) \), the limit of the statistic is always 1, and its distribution is

\[
N \left( 1 - S_1^{(1,2)}(x) \right) \cong \frac{1}{4} \xi_{11}^2 + \frac{1}{4} \xi_{22}^2 - \frac{1}{2} \xi_{11} \xi_{22} + \xi_{12}^2 = \left( \frac{\xi_{11} - \xi_{22}}{2} \right)^2 + \xi_{12}^2 = Q
\]

\[
\left( \begin{array}{cc}
\xi_{11} & \xi_{12} \\
\xi_{12} & \xi_{22}
\end{array} \right) \in \mathcal{N}(0, \Sigma)
\]

with \( \xi_1 = \xi_{11} - \xi_{22} \) and \( \xi_2 = \xi_{12} \), the rank of \( Q \) is 2, the matrix \( M \) of its coefficients is the identity and the covariance matrix of \([\xi_1, \xi_2]^T\) is

\[
\Sigma' = \left( \begin{array}{cc}
\frac{\sigma_{ii}^{ij} + \sigma_{jj}^{ij} - 2\sigma_{ij}^{ij}}{4} & \frac{\sigma_{ii}^{ij} - \sigma_{jj}^{ij}}{2} \\
\sigma_{ij}^{ij} & \sigma_{ij}^{ij}
\end{array} \right)
\]

where the values of \( \sigma_{ij}^{ij} \) are given in the proof of theorem 2. The characteristic function of \( Q \) has the form

\[
\phi(z) = \frac{1}{\sqrt{1 - 2iz\lambda_1 - 4z^2\lambda_2 - 2iz\lambda_2}}
\]

\( \lambda_1 \) and \( \lambda_2 \) are the eigenvalues of \( \Sigma' \).

Numerical example

Let us consider the most common case of \( S_1^{(1,2)} \). The above equations yield to

\[
\sigma_{ii}^{ij} = 2 \quad \sigma_{jj}^{ij} = 14 \quad \sigma_{ij}^{ij} = 4 \quad \sigma_{ii}^{ii} = 0 \quad \sigma_{jj}^{jj} = 0 \quad \sigma_{ij}^{ij} = 5
\]

\[
N \left( 1 - S_1^{(1,2)} \right) \in \phi(z) = \frac{1}{\sqrt{1 - 4iz\sqrt{1 - 10iz}}}\mu_{H_0} = 7 \quad \sigma_{H_0}^2 = 58
\]

We have drawn on figure 1 the Rice approximation of the asymptotic distribution and the histogram of 100000 trials of \( N \left( 1 - S_1^{(1,2)} \right) \) for large \( N \) (\( N = 10000 \)) which has the estimated first two cumulants : \( \mu_{H_0} = 6.997 \) and \( \sigma_{H_0}^2 = 57.63 \).

4.2 Distribution under the uniform hypothesis

We will give in this paragraph the limit distribution of the Hermite normality test under the nonnull particular hypothesis of normalised uniform \( H_1 \) : \( x \in \mathcal{U}([-\sqrt{3}, \sqrt{3}] \). In this case, the integral relation concerning Hermite polynomials

\[
\int_0^x H_n(t)dt = \frac{H_{n+1}(x) - H_{n+1}(0)}{n + 1}
\]

yield to the result for all \( i \)

\[
\mathbb{E} [\tilde{H}_i(x)] = \frac{H_{i+1}(\sqrt{3}) - H_{i+1}(-\sqrt{3})}{2\sqrt{3}(i + 1)!}
\]
which can be simplified to

\[ E \left[ \hat{H}_{2i+1}(x) \right] = 0 \quad \text{and} \quad E \left[ \hat{H}_{2i}(x) \right] = \frac{H_{2i+1}(\sqrt{3})}{\sqrt{3}(2i+1)(2i)!} \]

It is therewith easy to derive the expression of \( A = [a_{ij}] \) (refer to the notations (7))

\[
a_{ii} = E \left[ \hat{H}_i^2(x) \right] = \frac{1}{i!} \sum_{k=0}^{i} k! \binom{i}{k}^2 E \left[ \hat{H}_{2(i-k)}(x) \right] \tag{9}
\]

\[
a_{ij} = E \left[ \hat{H}_i(x)\hat{H}_j(x) \right] = \frac{1}{\sqrt{H_i}} \sum_{k=0}^{i} k! \binom{i}{k} \binom{j}{k} E \left[ \hat{H}_{i+j-2k}(x) \right] \tag{10}
\]

and these two equations allow us to obtain the values of \( h(A) \) and \( h'(A) \):

\[
h(A) = \frac{a_{ii}a_{jj} - (a_{ij})^2}{\left( a_{ii} + a_{jj} \right)^2}
\]

\[
h'(A) = \begin{pmatrix}
\frac{(a_{ij})^2-a_{ii}a_{jj}}{2} + (a_{ij})^2 & \frac{-a_{ij}}{a_{ii}+a_{jj}} \\
\frac{(a_{ij})^2-a_{ii}a_{jj}}{2} & \frac{(a_{ij})^2-a_{ii}a_{jj}}{2} + (a_{ij})^2
\end{pmatrix}
\]

Then, the limit distribution of the Hermite normality test is

\[
\sqrt{N} \left( S^{(i,j)}_H(x) - h(A) \right) \sim h'(A) \otimes \xi = h'_{11}\xi_{11} + h'_{22}\xi_{22} + 2h'_{12}\xi_{12}
\]
whose covariance matrix

\[
\Sigma = \begin{pmatrix}
\sigma_{i1}^{ij} & \sigma_{i2}^{ij} & \sigma_{i3}^{ij} \\
\sigma_{i1}^{ij} & \sigma_{i2}^{ij} & \sigma_{i3}^{ij} \\
\sigma_{i1}^{ij} & \sigma_{i2}^{ij} & \sigma_{i3}^{ij}
\end{pmatrix}
\]

is given by the equation (28).

\[
\sqrt{N} \left( S_H^{(i,j)}(x) - h(A) \right) \text{ is then asymptotically centered normal, with variance}
\]

\[
\sigma_{\mu H_1}^2 = (h_1' h_2' 2h_{12}') \Sigma \begin{pmatrix} h_{11}' \\
h_{12}' \\
2h_{12}'
\end{pmatrix}
\]

(11)

**Numerical example**

keeping the example of \( S_H^{(1,2)} \), we obtain

\[
\mu_{H_1} = A = \begin{pmatrix} 1 \\ 0.4 \\ 0 \end{pmatrix}, \quad h(A) = 0.8163
\]

\[
h'(A) = \begin{pmatrix} -0.35 \\ 0 \\ 0.8746 \end{pmatrix}
\]

\[
\Sigma = \begin{pmatrix} 0.8 & 0.2286 & 0 \\ 0.2286 & 0.1829 & 0 \\ 0 & 0 & 0.6286 \end{pmatrix}, \quad \sigma_{\mu H_1}^2 = 0.0979
\]

The figure 2 draws the histogram of 100000 trials of \( \sqrt{N} \left( S_H^{(1,2)}(x) - 0.8163 \right) \) for \( N = 10000 \) and the normal density \( N(0, 0.0979) \). The empirical mean and variance are \( \{ \hat{\mu}_{H_1} = 0.8162, \hat{\sigma}_{\mu H_1}^2 = 0.098 \} \).

### 4.3 On the choice of a best Hermite Normality test

As already told, the main advantage of the Hermite Normality test is that it defines a complete class of statistics, with two degrees of freedom: the number \( p \) of polynomials and their degrees. But how must we drive this choice in a practical application with real processes? There is obviously no real answer to this question because it depends highly on the data’s characteristics. Some tries have been made to solve this problem as well as possible. One example is a paper from R.V. Hogg [2] who gave a family of 4 test statistics based on the Kurtosis to test a density among a family of exponential distributions against another member of the same family. The most powerful statistic that tests the hypothesis of a fixed density against another fixed one are designed by Monte Carlo studies. Our approach determines theoretically the best test among the Hermite family against a fixed alternative - rather than by Monte Carlo means. The same method can be followed for any combination of polynomials. Some simple rules can be derived if we have some prior knowledge and by taking a look at the forms of Hermite polynomials. For instance, because a polynomial of degree \( d \) collects only powers with same parity as \( d \), while we consider a test with only odd degrees or only even degrees, the covariance matrix \( R \) will only depend on even order moments of the distribution and therefore, the test will better match the symmetrical alternatives. On the contrary, a test that cross the parity order of its Hermite polynomials will work for skewed alternatives.

Although we know usually nothing about the populations we have to test, we present in this section the means to determine an asymptotically best Hermite Normality test for a fixed alternative. Following the approach of section 4.2, we can find the limit distribution of the Hermite statistic under a fixed alternative distribution knowing its cumulants (moments). For that purpose, we consider two kinds of nonnormal distributions:

1. the beta distribution \( \beta(a, b) \)

\[
f(x) = \frac{x^{a-1}(1-x)^{b-1}}{\beta(a, b)} \quad x \in [0, 1]
\]
Asymptotic distribution of $S_H^{1,2}$ under the Uniform alternative distribution

$$
\mu_p = \mathbb{E}[x^p] = \frac{\beta(a + p, b)}{\beta(a, b)}
$$

2. The gamma distribution $\gamma(\alpha, \lambda)$

$$
f(x) = \frac{\alpha^\lambda x^{\lambda-1} e^{-\alpha x}}{\Gamma(\lambda)} \quad x \geq 0
$$

$$
\mu_p = \mathbb{E}[x^p] = \frac{\alpha^p \Gamma(\lambda + p)}{\Gamma(\lambda)}
$$

And using the general expression of Hermite polynomials, the relation

$$
\mathbb{E}[H_i(x)] = i! \sum_{k=0}^{[\frac{i}{2}]} \frac{(-1)^k \mu_{i-2k}}{2^k k!(i-2k)!} \quad (12)
$$

where $\mu_k$ are the standardised moments of the nonnormal distributions allows the calculation of $\mu_{H_1}$ and $\sigma_{H_1}^2$. This will ensure that we have access to the expressions of the asymptotic p.d.f. of $S_H$ under both the null and a nonnull particular hypothesis. We can then choose a best Hermite test with two polynomials using a classical Neyman-Pearson type decision, that is maximizing a correct detection probability (CDP) with a fixed false alarm probability (FAP).

$$
\text{CDP} = \text{Prob} \{ \text{choose } H_1 | H_1 \} = \text{Prob} \left\{ S_H < \eta \bigg| \sqrt{N}(S_H - h(A)) \in \mathcal{N}(0, \sigma_0^2) \right\} \quad (13)
$$

$$
\text{FAP} = \text{Prob} \{ \text{choose } H_1 | H_0 \} = \text{Prob} \{ S_H < \eta \big| N(1 - S_H) \in \text{Gamma}(\mu_{H_0}, \sigma_{H_0}^2) \} \quad (14)
$$

Because this kind of approach requires a lot of calculations, we have decided to take into account only the Hermite tests with two polynomials with orders less than 6. Table 1 gives three tests asymptotically the most powerful for a few nonnormal distributions. Because of the limits of this work (degrees < 6, only two polynomials,
asymptotic study), these results can not be taken as real rules, but rather as pieces of advice to choose a default test when nothing is known about the data. Actually, we have used these test to tabulate the quantiles in section 6.1.

<table>
<thead>
<tr>
<th>distribution</th>
<th>best three tests</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\beta(0.5, 0.5)$</td>
<td>$S_{H}^{(1,2)}$, $S_{H}^{(1,3)}$, $S_{H}^{(1,6)}$</td>
</tr>
<tr>
<td>$\beta(1.1)$ : uniform</td>
<td>$S_{H}^{(1,3)}$, $S_{H}^{(2,4)}$, $S_{H}^{(4,6)}$</td>
</tr>
<tr>
<td>$\beta(2, 2)$</td>
<td>$S_{H}^{(1,3)}$, $S_{H}^{(2,4)}$, $S_{H}^{(3,6)}$</td>
</tr>
<tr>
<td>$\beta(2, 3)$</td>
<td>$S_{H}^{(1,2)}$, $S_{H}^{(2,4)}$, $S_{H}^{(1,6)}$</td>
</tr>
<tr>
<td>Laplace</td>
<td>$S_{H}^{(1,2)}$, $S_{H}^{(1,6)}$, $S_{H}^{(2,6)}$</td>
</tr>
<tr>
<td>$\gamma(0.5, 1)$ : $\chi_{2}(2)$</td>
<td>$S_{H}^{(1,2)}$, $S_{H}^{(1,6)}$, $S_{H}^{(2,3)}$</td>
</tr>
<tr>
<td>$\gamma(0.5, 5)$ : $\chi_{2}(10)$</td>
<td>$S_{H}^{(1,2)}$, $S_{H}^{(1,6)}$, $S_{H}^{(2,4)}$</td>
</tr>
</tbody>
</table>

Table 1. most powerful tests for some alternatives

We moreover propose to verify that, for instance, $S_{H}^{(1,3)}$ is the most powerful test for the uniform hypothesis even for small sample sizes. We have then compared the empirical powers of the three tests given in table 1 with three other tests arbitrary chosen. Table 2 shows in percentage the powers (correct detection probability) for 6 Hermite tests given a 5% false alarm probability for sample sizes $N = 20, 30, 50$.

<table>
<thead>
<tr>
<th>$N$</th>
<th>$S_{H}^{(1,3)}$</th>
<th>$S_{H}^{(2,4)}$</th>
<th>$S_{H}^{(4,6)}$</th>
<th>$S_{H}^{(1,2)}$</th>
<th>$S_{H}^{(2,3)}$</th>
<th>$S_{H}^{(1,8)}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>20</td>
<td>0.57</td>
<td>0.63</td>
<td>0.23</td>
<td>0.19</td>
<td>0.03</td>
<td>0.04</td>
</tr>
<tr>
<td>30</td>
<td>0.79</td>
<td>0.16</td>
<td>0.39</td>
<td>0.32</td>
<td>0.07</td>
<td>0.08</td>
</tr>
<tr>
<td>50</td>
<td>0.96</td>
<td>0.64</td>
<td>0.67</td>
<td>0.59</td>
<td>0.08</td>
<td>0.27</td>
</tr>
</tbody>
</table>

Table 2. empirical powers of 6 Hermite tests for a normalized Uniform alternative

We can remark that these powers follow the asymptotic study especially since the size of the samples become larger. Actually, $S_{H}^{(1,3)}$ is the best test for all $N$, but $S_{H}^{(2,4)}$ and $S_{H}^{(4,6)}$ which are good for $N = 50$, work rather badly for smaller sizes. This emphasizes the fact that we have to take care when applying asymptotic results to small samples ($N < 100$). In particular, when we hesitate between some tests (reason why we gave three tests in table 1), the simpler but intuitive rule when the length of the samples decreases is to choose the test with the polynomials of smaller degrees, which are clearly more robust.

5 Exact distribution for finite $N$

In order to complete the asymptotic study, it is of a great interest to have access to the exact distribution of a test statistic for all sample sizes (and especially for small ones). The exact distribution of $S_{H}$ can not yet be found with limit theorems and the structure of the statistic seems not to provide easily usable statistical results. By the way, a solution is to approximate the empirical distribution of $S_{H}$ by an appropriate one with explicit p.d.f.

The compromise between the number of parameters which define the model and the relevancy of the fitting is important. In our case, depite the interest of explaining some theoretical behaviour of the Hermite test when the polynomials and the sample size are changing, the real aim is to be able to calculate the quantiles of every order with a unique model (see section 6.1). We present briefly thereafter two different means of fitting the empirical distribution of $S_{H}$, having always in mind that our main concern is to reduce the number of parameters.

Remark: location and scale invariance

In the sequel of the article, we consider samples with unknown mean and variance, which makes the Hermite test scale and location invariant. The property of orthogonality between Hermite polynomials which is the basis of the Hermite test so requires a standardisation of the populations with their sample mean and variance: $\bar{x} = \frac{x - \mu}{\sigma}$. The modelisation of $S_{H}$ for finite $N$, the calculation of the quantiles and the comparison of power have been conducted with this modification of the test.

11
The reason why we introduce the standardisation at this point of the study is that it affects the asymptotical behaviour. Actually, the limit distribution is still a quadratic form under $H_0$ and normal under $H_1$, because those results rely on the structure of the sphericity statistic, but their parameters cannot be derived as easy as in the standardised case. Those considerations are fully detailed in [14].

\section{Fitting by a Pearson curve}

Mauchly [7] approximates the nonasymptotic distribution of the sphericity test by a beta distribution with parameters depending on the sample moments. This fact and the general remarks about the approximation of empirical distributions in [15] led us to fit the exact distribution of $S_H$ by the Pearson curves family. The choice of the distribution is driven by the values of the normalized sample skewness and kurtosis [13]. Lots of simulations provide either a Pearson type I (beta) or a Pearson type III (gamma) distribution. Because our aim is in fact to approximate accurately the tails of the distribution (to calculate the quantiles) and the beta distribution gives better results for the tails, we have decided to fit the data only with a type I Pearson curve with same mean and variance:

$$
g(x) = \frac{x^{a-1}(1-x)^{b-1}}{\beta(a,b)}
$$

with

$$
m = \frac{\beta(a+1,b)}{\beta(a,b)} \quad \sigma^2 = \frac{\beta(a+2,b)}{\beta(a,b)} - \left(\frac{\beta(a+1,b)}{\beta(a,b)}\right)^2
$$

If this kind of distribution work good for some Hermite tests like $S_H^{(1,2)}$, it becomes very bad for higher order polynomials. The figures 3-4 draw the shapes of the empirical distribution (solid line) and the corresponding beta curve (with "o") for $N = 50$ and the two tests $S_H^{(1,2)}$ and $S_H^{(1,4)}$. We need then to improve the quality of the approximation, what we will do in the next section.

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{figure3.png}
\caption{Exact distribution of $S_H^{(1,2)}$ and the corresponding beta model}
\end{figure}
5.2 Mixture models

We propose in this section an approach based on the mixture of two distributions. We have considered two kinds of mixtures: the very classical model with normal p.d.f., and a mixture of two beta distributions regarding the results of the preceding section. The density functions are also depending on 5 parameters:

\[
g_{\text{mixt}, \mathcal{N}}(x) = \alpha \mathcal{N}(\mu_1, \sigma_1^2) + (1 - \alpha) \mathcal{N}(\mu_2, \sigma_2^2)
\]

\[
g_{\text{mixt}, \beta}(x) = \alpha \beta(a_1, b_1) + (1 - \alpha) \beta(a_2, b_2)
\]

The identification of the parameters are achieved with a classical Gauss-Newton algorithm that tends to minimize a cost function under some constraints. We have chosen a quadratic cost function because it exhibits almost the same performances as with a Kullback divergence, but is simpler to implement and the constraints are build with the first four moments.

\[
J_{mc} = \sum_{i \in I} \left( \hat{f}(\bar{x}_i) - g_{\text{mixt}}(\bar{x}_i) \right)^2
\]

where \( \hat{f}(\bar{x}_i) \) is the histogram estimated at the points \( \bar{x}_i \) in \( I = [0, 1] \).

In order to decide which model (between the beta distribution and the two mixtures) best fits the empirical distribution of \( S_H^{(1,4)} \) for a set of Hermite polynomials and a fixed \( N \), we introduce a simple normalized quadratic distance

\[
D(\hat{f}, g) = \sqrt{\sum_{i \in I} \left( \frac{\hat{f}(\bar{x}_i) - g(\bar{x}_i)}{\hat{f}(\bar{x}_i)} \right)^2}
\]

To illustrate the gain obtained with the use of mixture model, we have drawn on figure 5 the distribution of \( S_H^{(1,4)} \) with its corresponding beta mixture. The parameters of the fitted distributions chosen with the procedure above described have been tabulated for a lot of Hermite tests and for sample sizes \( N \in \{20, 30, 40, 50, 75, 100, 150, 200\} \). For obvious reasons of size, we can not present these tables here and they are reported in [14].
Figure 5. Exact distribution of $S_H^{(1,4)}$ and the corresponding beta-mixture model

6 Simulations and Monte Carlo study

6.1 Quantiles of the Hermite Normality test

The power of a test statistic of normality can be defined in different ways, depending on the correct detection and the false alarm probability under both hypothesis. We follow the work of Pearson and al. [3] which consider a wide range of nonnormal populations and give the correct detection probability under $H_1$ with a fixed false alarm probability under $H_0$. The quantile of order $\alpha$ of $S_H$ is the threshold (within $[0, 1]$) that provides a $100\%$ false alarm probability under the normal hypothesis.

$$\int_0^\alpha f_0(\bar{x})d\bar{x} = \alpha$$

with $f_0(\bar{x})$ the density of $S_H$ under $H_0$.

We can not calculate these quantiles theoretically because we do not have a close form for the probability density function $f_0(\bar{x})$. We can therefore use the parametrisation made in section 5 or estimate them empirically. For that purpose, we have considered a large number of trials (500000) of $S_H$ under $H_0$ with a sample size $N$. The quantiles of order $\alpha \in \{0.01, 0.05, 0.1\}$ have been tabulated for 12 Hermite tests choosen with the help of table 1 and for different samples sizes $N \in \{20, 30, 40, 50, 75, 100, 150, 200\}$ (see tables 9-11).

6.2 Comparison of power

A lot of works have been achieved about the comparison of power between tests of normality since few decades [1] [3] [4]. We have followed the comments already done within to compare the Hermite test only to three other normality tests which are based on radically different structures.

(i) The Anderson-Darling test $A^2$ that make use of the empirical distribution function (EDF) ([16], pp. 372),

(ii) the D’Agostino $D$ test based on order statistics [17],
and a test with sample moments of order 3 and 4 due to D’Agostino and Pearson [18].

We could have replaced $A^2$ by the famous Shapiro-Wilkes $W$ test, but Dyer [19] says that they have similar powers. The approach we have followed in the same as in [3] and has three steps:

1. first, we draw 5000 samples of size $N \in \{20, 50, 100\}$ from a wide range of nonnormal distributions given in table 5. Note that we must use standard forms of those distributions.

2. the second step is to apply these populations to the three tests above described and some Hermite Normality tests,

3. then, we record each time the value of the statistic lie within the quantiles (under or above depending on the test) and dividing this count by 50 gives immediately an estimation (in %) of the power of a test for a fixed $N$ and a quantile of order $\alpha$.

The same kind of work has already been done for some Hermite tests with two polynomials [5]. In order to show different and complementary results, we have made the experiments with 4 Hermite tests with 3 polynomials and one with 4 polynomials. The tests presented here have been chosen through simulations over a large number of polynomial sets and by considering the ones which were overall the most powerful. We need also the values of the quantiles of these tests and the three alternative tests considered. They are drawn in tables 3 and 4.

<table>
<thead>
<tr>
<th>$N$</th>
<th>$A^2$</th>
<th>$R^2$</th>
<th>$D$</th>
</tr>
</thead>
<tbody>
<tr>
<td>20</td>
<td>0.753</td>
<td>6.712</td>
<td>$[-3.958 \quad -0.458]$</td>
</tr>
<tr>
<td>50</td>
<td>0.741</td>
<td>6.475</td>
<td>$[-3.307 \quad 0.359]$</td>
</tr>
<tr>
<td>100</td>
<td>0.738</td>
<td>6.281</td>
<td>$[-2.934 \quad 0.807]$</td>
</tr>
</tbody>
</table>

Table 3. Quantiles of order 0.05 for the alternative tests

<table>
<thead>
<tr>
<th>$N$</th>
<th>$S^{(1,2,3)}_H$</th>
<th>$S^{(1,2,6)}_H$</th>
<th>$S^{(1,3,6)}_H$</th>
<th>$S^{(1,3,9)}_H$</th>
<th>$S^{(1,2,3,4)}_H$</th>
</tr>
</thead>
<tbody>
<tr>
<td>20</td>
<td>0.158</td>
<td>0.232</td>
<td>0.162</td>
<td>0.0462</td>
<td>0.02</td>
</tr>
<tr>
<td>50</td>
<td>0.322</td>
<td>0.411</td>
<td>0.307</td>
<td>0.186</td>
<td>0.054</td>
</tr>
<tr>
<td>100</td>
<td>0.475</td>
<td>0.541</td>
<td>0.409</td>
<td>0.327</td>
<td>0.107</td>
</tr>
</tbody>
</table>

Table 4. Quantiles of order 0.05 for 5 Hermite tests

19 populations have been drawn from the distributions of table 5. The ones considered far from the normal have been tested with few samples $N = 20$, the ones closer to the normal with $N = 100$ and the rest with the intermediate size $N = 50$. The powers are pointed out in tables 6-8. The values of the third and fourth order standardised moments are given to indicate how far from the normal distribution the populations are.

\[
\begin{align*}
\sqrt{\beta_1} &= \sqrt{\mu_3^2} \\
\beta_2 &= \mu_4
\end{align*}
\]

(20)

To ease the analysis of the results, we have emphasized in each case the most powerful test in bold.

We need obviously to cross the results of the tables 6-8 with those in [5] which consider 5 other Hermite tests. A common remark is that the Hermite tests seem to work best for platykurtic populations than for leptokurtic ones.

Moreover, the use of $H_1(x)$ in the construction of a test is decisive (that is why we consider here this kind of tests) because the information carried by the variance of a sample is important and contributes to the robustness of the Hermite statistic. Finally, when gathered, the powers of the Hermite Normality tests are the greatest for 10 populations amongst the 19 tested (12 if we consider $S^{(1,2)}_H$ too). This ensures that additionally to the modular structure of the Hermite test, it works rather well compared to the other normality tests proposed in the litterature.

If one has to test a sample without any prior knowledge, we can advise the use of $S^{(1,2)}_H$, $S^{(1,3)}_H$ and $S^{(1,2,3)}_H$ which were the more robusts for the nonnormal populations studied, but having always in mind that another set of polynomials could provide better results.
distribution | equation
--- | ---
Beta | \( f(x) = \frac{\beta(a, b)}{x^a / (x/b_n z^r)} = x^{a/b_n z^r} \)
Gamma | \( f(x) = \frac{\Gamma(p)}{\Gamma(e^{\beta x})} = x^{p / (1 + e^{\beta x})^{p+1}} \)
Student’s t | \( f(x) = \frac{\Gamma(\frac{n}{2})}{\sqrt{n\pi(1 + \frac{x^2}{\nu})^{p+1}}} \)
Weibull | \( f(x) = k x^{k-1} e^{-x^k} \)
Logistic | \( f(x) = \frac{\beta e^{\beta x}}{(1 + e^{\beta x})^2} \)
Laplace | \( f(x) = \frac{1}{2} e^{-|x|} \)
Johnson’s SB | \( X = \gamma + \delta \log \frac{x}{1-x} \)
Johnson’s SU | \( X = \gamma + \delta \sinh^{-1}(y) \)
Lognormal | \( X = \gamma + \delta \log(1 - x) \)
Tukey | \( x = R^\lambda - (1 - R)^\lambda \)

| population | \( \sqrt{\beta_1} \) | \( \beta_2 \) | \( A^2 \) | \( K^2 \) | \( D \) | \( S_H^{(1,2,3)} \) | \( S_H^{(1,2,6)} \) | \( S_H^{(1,3,3)} \) | \( S_H^{(1,3,9)} \) | \( S_H^{(1,2,3,4)} \)
--- | --- | --- | --- | --- | --- | --- | --- | --- | --- | ---
symmetrical | | | | | | | | | |
beta(0.5, 0.5) | 0 | 1.5 | 68 | 45 | 4 | 88 | 37 | 78 | 83 | 75
 SB(\( \gamma = 0, \delta = 0.5 \)) | 0.63 | 42 | 27 | 7 | 71 | 25 | 61 | 71 | 50
t(\( \nu = 1 \)) | 0 | \( \times \) | 87 | 72 | 89 | 74 | 71 | 76 | 66 | 74
skewed | | | | | | | | | |
\( \chi^2 (\nu = 2) \) | 2 | 9 | \( \times \) | 79 | 48 | 52 | 63 | 60 | 41 | 21 | 46
lognormal | 6.2 | 114 | \( \times \) | 89 | 71 | 88 | 74 | 69 | 77 | 66 | 74
Cauchy | | | | | | | | | |

Table 5. Nonnormal distributions

Table 6. Estimated powers : \( \alpha = 0.05 \) \( N = 20 \)

7 Conclusion and perspectives

We have made in this paper a large study of the Hermite Normality test previously reported in [5]. The asymptotic distribution has been found both for the null and the nonnull hypothesis. The fact that the distribution under a fixed alternative depends on the parameters of the Hermite test led us to try to find the one asymptotically most powerful among some fixed alternatives. In the case of finite sample sizes, the distribution of \( S_H \) has been approximated by either a beta curve, either a mixture model. The quantiles of order \( \alpha \in \{0.01, 0.05, 0.1\} \) have been tabulated for 12 Hermite tests and different samples sizes, and a comparison of power with three other tests has been made. The results show that our test works well for a lot of nonnormal populations and especially for platykurtic ones. This reinforces the main advantage of the Hermite Normality test which is to define a complete class of test statistic with the same structure but with different behaviours. The quality of the results obtained drive us to extend the Hermite test to correlated data or to multivariate populations with the help of some powerful formulas on Hermite polynomials of several variables. This will be the purpose of a future work.

ACKNOWLEDGEMENTS

We wish to thank Ghada Jammal to have started the exploration of this normality test and therefore given the first evidences of its efficiency. We also thank the referees for their helpfull remarks.
\[
\begin{array}{|c|c|c|c|c|c|c|c|c|c|}
\hline
\text{population} & \sqrt{\beta_1} & \beta_2 & A^2 & K^2 & D & S_{H}^{[1,2,3]} & S_{H}^{[1,2,6]} & S_{H}^{[1,3,6]} & S_{H}^{[1,3,9]} & S_{H}^{[1,2,3,4]} \\
\hline
\text{symmetrical} & & & & & & & & & & \\
\text{uniform} (\beta(1,1)) & 0 & 1.8 & 63 & 76 & 58 & 90 & 93 & 87 & 94 & 91 \\
Tukey (\lambda = 2.5) & 0 & 1.9 & 41 & 56 & 55 & 76 & 83 & 75 & 87 & 77 \\
Laplace & 0 & 6 & 53 & 47 & 61 & 35 & 33 & 40 & 30 & 33 \\
t (\nu = 4) & 0 & 14 & 40 & 45 & 52 & 43 & 41 & 46 & 38 & 42 \\
t (\nu = 2) & 0 & \times & 86 & 76 & 90 & 78 & 75 & 81 & 75 & 75 \\
\hline
\text{skewed} & & & & & & & & & & \\
S_{H} (\gamma = 1, \delta = 1) & 0.7 & 2.9 & 71 & 37 & 11 & 71 & 34 & 13 & 7 & 24 \\
\chi^2 (\nu = 10) & 0.9 & 4.2 & 49 & 44 & 27 & 54 & 41 & 34 & 20 & 39 \\
Erlang (\nu = 2) & 1.42 & 6 & 91 & 67 & 58 & 92 & 74 & 57 & 35 & 67 \\
\hline
\end{array}
\]

Table 7. Estimated powers : \( \alpha = 0.05 \) \( N = 50 \)

<table>
<thead>
<tr>
<th>population</th>
<th>( \sqrt{\beta_1} )</th>
<th>( \beta_2 )</th>
<th>( A^2 )</th>
<th>( K^2 )</th>
<th>( D )</th>
<th>( S_{H}^{[1,2,3]} )</th>
<th>( S_{H}^{[1,2,6]} )</th>
<th>( S_{H}^{[1,3,6]} )</th>
<th>( S_{H}^{[1,3,9]} )</th>
<th>( S_{H}^{[1,2,3,4]} )</th>
</tr>
</thead>
</table>
| symmetrical | & & & & & & & & & & \\
| \( \beta(2,2) \) & 0 & 2.1 & 36 & 62 & 71 & 59 & 68 & 63 & 79 & 68 \\
| \( S_{U} (\gamma = 0, \delta = 3) \) & 0 & 3.5 & 10 & 17 & 16 & 15 & 13 & 16 & 13 & 14 \\
| logistic & 0 & 4.2 & 23 & 33 & 38 & 28 & 25 & 31 & 27 & 27 \\
| skewed | & & & & & & & & & & \\
| \( \beta(2,3) \) & 0.3 & 2.4 & 43 & 38 & 30 & 51 & 32 & 20 & 25 & 37 \\
| \( Weibull (k = 2) \) & 0.6 & 3.25 & 61 & 52 & 15 & 76 & 45 & 25 & 16 & 50 \\
| \hline

Table 8. Estimated powers : \( \alpha = 0.05 \) \( N = 100 \)

A Statistical properties of Hermite polynomials

To have a full description of Hermite polynomials (generation, properties, series expansions ...), one can refer to [20][21].

- Let \( x \) be a random variate normally distributed with zero mean and unit variance. With some combinatorial relations and the properties of Hermite polynomials, we can express the expectation of products of \( H_i(x) \). Without loss of generality, we assume that the polynomials are sorted in ascending order, that is \( m \leq n \leq p \leq q \):

\[
\mathbb{E} [H_n(x)] = 0 \tag{21}
\]

\[
\mathbb{E} [H_m(x)H_n(x)] = n! \delta(n-m) \tag{22}
\]

\[
\mathbb{E} [H_m(x)H_n(x)H_p(x)] = \binom{m}{k} \binom{n}{k} k!p! \quad \text{with} \quad \left\{ \begin{array}{l}
\frac{k}{p} = \frac{m+n-k}{p} \\
\frac{q}{p} = \frac{m+n}{p}
\end{array} \right. \tag{23}
\]

from which we deduce

\[
\mathbb{E} [H_{2n+1}^2(x)] = 0 \quad \mathbb{E} [H_{2n}^2(x)] = \left( \frac{2n}{n!} \right)^2 \tag{24}
\]

\[
\mathbb{E} [H_m(x)H_n(x)H_p(x)H_q(x)] = \sum_{k=0}^{\min(m, \frac{n+k-p+1}{2})} k!k!(m+n-2k)!(\frac{m}{k})!\binom{n}{k}!\binom{q}{k'}! \quad \text{with} \quad \left\{ \begin{array}{l}
k' = k + \frac{p+q-m-n}{2} \\
q' = m+n+p
\end{array} \right. \tag{25}
\]

which drives to

\[
\mathbb{E} [H_m^2(x)H_n^2(x)] = m!n! \sum_{k=0}^{m} \binom{m}{k} \binom{n}{k} \binom{2k}{k} \tag{26}
\]
Table 9. Quantiles of the Hermite Normality test: $\alpha = 0.01$

Table 10. Quantiles of the Hermite Normality test: $\alpha = 0.05$

- $x$ is now a random variate with unknown distribution $f(x)$. Using only combinatorial expansions of Hermite polynomials, one can derive the same kind of relations than (22) and (25):

\[
\mathbb{E} [H_m(x)H_n(x)] = \sum_{k=0}^{m} k! \binom{m}{k} \binom{n}{k} \mathbb{E} [H_{m+n-2k}(x)]
\]

\[
\mathbb{E} [H_m(x)H_n(x)H_p(x)H_q(x)] = \sum_{k=0}^{m} \sum_{k'=0}^{p} k! k'! \binom{m}{k} \binom{n}{k} \binom{p}{k'} \binom{q}{k'} \min\{m+n-2k, p+q-2k'\} \mathbb{E} [H_{m+n+p+q-2(k+k'+k'')} (x)]
\]

References


<table>
<thead>
<tr>
<th>N</th>
<th>$S_{(1,2)}^H$</th>
<th>$S_{(1,3)}^H$</th>
<th>$S_{(1,4)}^H$</th>
<th>$S_{(1,5)}^H$</th>
<th>$S_{(2,3)}^H$</th>
<th>$S_{(2,4)}^H$</th>
<th>$S_{(2,6)}^H$</th>
<th>$S_{(3,5)}^H$</th>
<th>$S_{(4,6)}^H$</th>
<th>$S_{(1,2,3)}^H$</th>
<th>$S_{(1,2,4)}^H$</th>
</tr>
</thead>
<tbody>
<tr>
<td>20</td>
<td>0.685</td>
<td>0.349</td>
<td>0.601</td>
<td>0.628</td>
<td>0.529</td>
<td>0.481</td>
<td>0.148</td>
<td>0.467</td>
<td>0.161</td>
<td>0.121</td>
<td>0.201</td>
</tr>
<tr>
<td>30</td>
<td>0.767</td>
<td>0.457</td>
<td>0.66</td>
<td>0.674</td>
<td>0.587</td>
<td>0.519</td>
<td>0.18</td>
<td>0.546</td>
<td>0.193</td>
<td>0.15</td>
<td>0.28</td>
</tr>
<tr>
<td>40</td>
<td>0.813</td>
<td>0.531</td>
<td>0.692</td>
<td>0.702</td>
<td>0.621</td>
<td>0.553</td>
<td>0.211</td>
<td>0.596</td>
<td>0.212</td>
<td>0.176</td>
<td>0.341</td>
</tr>
<tr>
<td>50</td>
<td>0.844</td>
<td>0.585</td>
<td>0.711</td>
<td>0.72</td>
<td>0.644</td>
<td>0.584</td>
<td>0.241</td>
<td>0.628</td>
<td>0.227</td>
<td>0.197</td>
<td>0.391</td>
</tr>
<tr>
<td>75</td>
<td>0.888</td>
<td>0.676</td>
<td>0.741</td>
<td>0.745</td>
<td>0.681</td>
<td>0.642</td>
<td>0.309</td>
<td>0.674</td>
<td>0.253</td>
<td>0.236</td>
<td>0.485</td>
</tr>
<tr>
<td>100</td>
<td>0.913</td>
<td>0.73</td>
<td>0.758</td>
<td>0.758</td>
<td>0.707</td>
<td>0.683</td>
<td>0.365</td>
<td>0.7</td>
<td>0.27</td>
<td>0.26</td>
<td>0.551</td>
</tr>
<tr>
<td>150</td>
<td>0.939</td>
<td>0.796</td>
<td>0.777</td>
<td>0.773</td>
<td>0.74</td>
<td>0.744</td>
<td>0.452</td>
<td>0.732</td>
<td>0.295</td>
<td>0.29</td>
<td>0.638</td>
</tr>
<tr>
<td>200</td>
<td>0.953</td>
<td>0.834</td>
<td>0.79</td>
<td>0.781</td>
<td>0.759</td>
<td>0.782</td>
<td>0.515</td>
<td>0.751</td>
<td>0.315</td>
<td>0.309</td>
<td>0.695</td>
</tr>
</tbody>
</table>

**Table 11.** Quantiles of the Hermite Normality test: $\alpha = 0.1$


