Developments in Statistics
Andrej Mrvar and Anuška Ferligoj (Editors)
Metodološki zvezki, 17, Ljubljana: FDV, 2002

Testing Symmetry by an Easy-to-Calculate Statistic Based on Letter Values

Maurizio Brizzi[1]

Abstract

This paper deals with an index of skewness (the letter coefficient of skewness, *l.c.s.*), proposed in a previous study (Brizzi, 2000), which is based on a particular set of sample quantiles called *letter values*, introduced by Tukey (1977). Some analytical properties of this index are described, and the sample distribution of this index is simulated under symmetric (normal and uniform) and positively skewed (exponential, Gamma, Log-normal and Skew-normal) distribution models. The power of a test based on the *l.c.s.* value is then compared with the "classic" test focused on Pearson third-moment skewness index.

1 Introduction

The problem of checking the shape (skewness and kurtosis) of a population (or set of data) has been challenging a lot of methodologists and applied statisticians. The traditional Pearson statistic $\sqrt{b_1}$ (index of skewness) and b_2 (index of kurtosis) based on moments are surely the most diffused indices of shape, but a lot of alternative indices have been proposed, using the sample information in different ways. Oja (1981) tries to define partial orderings of distributions by means of descriptive indices, including indices of shape; Mac Gillivray (1986) gives a thorough classification of the measures of shape. The works of Ruppert (1987), Balanda and Mac Gillivray (1988), Zenga (1996) deal with the real meaning of kurtosis. Many Authors, such as Tukey (1977), Hoaglin (1985), Kappenman (1988), Moors (1988), Groeneveld (1998), Brizzi (2000) have focused their analysis on quantiles and order statistics, as worth tools for evaluating the degree

[1] Dipartimento di Scienze statistiche, Via delle Belle Arti 41, I-40126 Bologna, Italy.

of skewness or kurtosis of a sample, and making inference about the corresponding population.

Let $x_1, x_2, \ldots, x_n$ be a simple random sample from a certain population P, and *m(x)* and *s(x)* be the sample mean and the sample standard deviation. Tukey (1977) defined and Hoaglin (1985) developed a particular set of order statistics, called the *letter values*, by considering:

- the sample median (Tukey uses M for median, but we prefer to use H for half, in order to avoid any confusion with the averaging operator)
- the lower and upper quartiles (F+ and F- for fourth)
- the lower and the upper octiles (E+ and E- for eighth), and so on, halving the extreme sets of data and going backwards with the alphabet: D, C, B, A, Z, Y etc.

The last letter values are the sample extremes; then, the number k_n of distinct pairs of letter values is a function of the sample size, and precisely:

$$kn = | \log2 n | + 1 \qquad (1.1)$$

For example, we have $k_{10} = 4$, $k_{40} = 6$. The average of each pair of corresponding letter values is called a *midsummary* (denoted by *mid-F, mid-E, mid-D,…* depending on the letter values involved). When the sample is perfectly symmetric, all the midsummaries are equal to the median; then, is it possible to use them for checking the skewness of a sample. Tukey and Hoaglin proposed a graphical study of the midsummaries; Brizzi (2000) defines a coefficient of skewness based on them. The main aim of this paper is to study, by Monte Carlo simulation, the behavior of such coefficient and the power of the related test against some particular alternatives.

2 The letter coefficient of skewness (L.C.S.)

The *letter coefficient of skewness* (*l.c.s.* for short), recently introduced by Brizzi (2000), is a measure of skewness based on letter values, very quick to compute even when the sample is large. Supposing that the sample size allows to define k_n distinct midsummaries, and denoting the *i*-th midsummary by m_i ($m_o = H$, $m_1 =$ *mid-F*, $m_2 =$ *mid-E*, …, up to m_k), the *l.c.s.* is defined as the least-squares estimate of the slope of the following linear model:

$$z_i = a + b\, t_i + e_i \qquad (2.1)$$

fitted to the set of k_n+1 points (t_i , z_i), where:

$$t_i = \frac{i}{k_n} \; ; \; i = 0, 1, ..., k_n \qquad (2.2)$$

$$z_i = \frac{m_i - m(x)}{s(x)} \;;\; i = 0, 1, ..., k_n \tag{2.3}$$

The error terms e_i in (2.1) are not to be considered i.i.d. as usually; the slope b is a purely descriptive statistic whose properties are yet to be explored.

It is not difficult to demonstrate (see Brizzi, 1999) that the coefficient b may be written as a linear combination of the standardized midsummaries z_i's:

$$b = \frac{12k_n}{(k_n+1)(k_n+2)} \sum_{i=1}^{k_n} z_i \left(t_i - \frac{1}{2} \right) \tag{2.4}$$

Example. If our data are the squares of the natural numbers from 1 to 21, e.g., $x_1 = 1$, $x_2=4$, …, $x_{21} = 441$, we have $k_{21} = 5$. The midsummaries are: H=121, mid-F=146, mid-E=177.5; mid-D=202; mid-C=211.5, mid-B = 221. The trend is increasing, denoting a positive skewness. The l.c.s. is equal to *0.7509.* The Pearson statistic $\sqrt{b_1}$, calculated on the same set of data, is equal to 0.6094.

The expression of b in (2.4) is a weighted sum of part of data; then, if the population has at least two finite moments, the Lindeberg-Feller Central Limit Theorem holds; therefore, the sample distribution of b is asymptotically normal. Moreover, it can be easily shown that the *l.c.s.* is invariant under increasing linear transformations: in fact, all the z_i's satisfy the same invariance property, due to the standardization (2.3). Then, if p is a positive number and q a real number, we have:

$$b(X) = b\,(pX + q) \tag{2.5}$$

Due to this property, if our data are lengths expressed in centimeters or inches, or temperatures expressed in degrees Celsius or Fahrenheit, the *l.c.s.* does not depend on the scale of measure used.

3 Sample distribution of the L.C.S. under symmetric models

The *l.c.s.* may be used for testing symmetry of the population from which the data are coming. Evidently, it is possible to define an exact sized test only when the sample distribution of the test statistic a is known (at least approximately) under the null hypothesis. We have studied here the distribuiton of the *l.c.s.* under normality, and we propose then a quick test of symmetry in a "nearly Gaussian" environment, supposing that the sample data are approximately unimodal and the tails are not too heavy. When these conditions hold, we need only symmetry for normality. With this aim, we have simulated, using the GAUSS package, for each sample size (twelve different sizes between 20 and 100), a fixed number of samples (30,000) of i.i.d. observations from a standard normal distribution. We have decided to consider only samples which can give at least five couples of letter values, but it would have been technically possible also to consider smaller samples. The invariance property (2.5) ensures us that for Gaussian variables, the sample distribution of the *l.c.s.* is the same; therefore, we have no loss in generality. In Table 1 we have reported the main features of sample distribution of the statistic. Looking at the table, we can observe that simulated sample distribution of *l.c.s.* is almost perfectly symmetric around zero: the corresponding tail percentiles are very near in absolute value, and the mean and median are very near to zero, even when sample size is small. The tail percentiles may be used as threshold values for the critical region of the statistic.

Table 1: Sample distribution of the l.c.s.under a N(0,1) distribution.

				Left tail percentiles				***Right tail percentiles***		
n	k_n	*Average*	*StD*	*1%*	*2.5%*	*5%*	*50%*	*95%*	*97.5 %*	*99 %*
20	*5*	*0.000*	*0.411*	*-0.936*	*-0.800*	*-0.675*	*-0.001*	*0.680*	*0.805*	*0.943*
25	*5*	*-0.004*	*0.401*	*-0.924*	*-0.790*	*-0.666*	*-0.005*	*0.654*	*0.778*	*0.919*
30	*5*	*-0.001*	*0.388*	*-0.898*	*-0.762*	*-0.638*	*0.001*	*0.639*	*0.757*	*0.900*
35	*6*	*-0.001*	*0.372*	*-0.868*	*-0.736*	*-0.613*	*-0.001*	*0.610*	*0.728*	*0.866*
40	*6*	*0.001*	*0.368*	*-0.855*	*-0.725*	*-0.609*	*0.000*	*0.603*	*0.721*	*0.854*
45	*6*	*0.002*	*0.362*	*-0.847*	*-0.710*	*-0.595*	*0.003*	*0.597*	*0.713*	*0.846*
50	*6*	*-0.003*	*0.351*	*-0.834*	*-0.691*	*-0.582*	*-0.003*	*0.573*	*0.687*	*0.821*
60	*6*	*0.000*	*0.343*	*-0.807*	*-0.673*	*-0.566*	*+0.002*	*0.566*	*0.676*	*0.805*
70	*7*	*0.001*	*0.333*	*-0.776*	*-0.657*	*-0.546*	*+0.002*	*0.549*	*0.658*	*0.787*
80	*7*	*-0.002*	*0.327*	*-0.772*	*-0.646*	*-0.539*	*-0.002*	*0.536*	*0.642*	*0.768*
90	*7*	*-0.002*	*0.323*	*-0.756*	*-0.637*	*-0.531*	*0.000*	*0.530*	*0.635*	*0.760*
100	*7*	*-0.002*	*0.313*	*-0.745*	*-0.621*	*-0.519*	*-0.002*	*0.509*	*0.615*	*0.737*

In Fig.1 we have represented the sample distribution, under normality, of the l.c.s. for three different values of n. Looking at the figure, we can come to the same conclusions: the sample distribution is symmetric with respect to zero and the dispersion slightly decreases when the sample size increases.

The symmetry of the sample distribution of the l.c.s. is quite evident; then, it has no sense to use different limit values (except for the sign) for the two tails. Then, we have rewritten the tail values of such distribution by taking the half sum of the corresponding tail values. The resulting tail percentiles, whose trend seems to be more regular than in Table 1, have been reported in Table 2, and may be used as critical values in a test of symmetry.

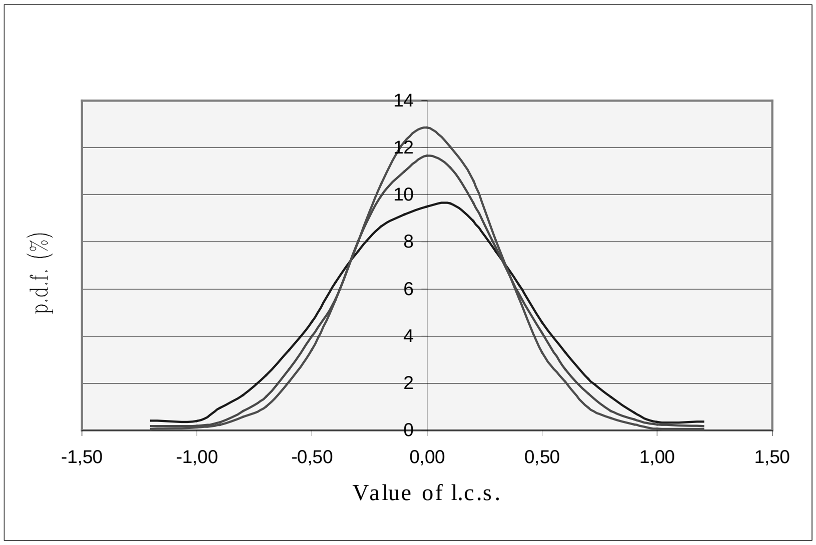

Figure 1: Sample distribution of the l.c.s. when sampling from a normal population (lower line: n=25; median line: n=50; upper line: n=100).

Table 2: Modified tail percentiles of the l.c.s.

		Right tail percentiles		
n	k_n	*95%*	*97.5%*	*99%*
20	*5*	*0.678*	*0.802*	*0.940*
25	*5*	*0.660*	*0.784*	*0.921*
30	*5*	*0.638*	*0.760*	*0.899*
35	*6*	*0.612*	*0.732*	*0.867*
40	*6*	*0.606*	*0.723*	*0.855*
45	*6*	*0.596*	*0.712*	*0.846*
50	*6*	*0.578*	*0.689*	*0.828*
60	*6*	*0.566*	*0.675*	*0.806*
70	*7*	*0.548*	*0.658*	*0.782*
80	*7*	*0.538*	*0.644*	*0.770*
90	*7*	*0.530*	*0.636*	*0.758*
100	*7*	*0.514*	*0.618*	*0.741*

We have also simulated the distribution of the l.c.s. under another symmetric model, with a markedly different shape: the uniform distribution. With no loss of generality we simulated uniformly distributed values from zero to one. The results are reported in Table 3.

The uniform-generated distribution is evidently symmetric and the mean-median value is zero. If we make a comparison between Table 1 and Table 3, we notice that the latter has less dispersion (indeed, the standard deviation takes lower values), therefore the tail percentiles are closer to zero.

Table 3: Sample distribution of the l.c.s.under an uniform U(0,1) distribution.

				Left tail percentiles				***Right tail percentiles***		
n	k_n	*Average*	*StD*	*1%*	*2.5%*	*5%*	*50%*	*95%*	*97.5%*	*99 %*
20	*5*	*0.0012*	*0.335*	*-0.764*	*-0.651*	*-0.548*	*-0.0007*	*0.553*	*0.656*	*0.763*
25	*5*	*0.0005*	*0.317*	*-0.726*	*-0.617*	*-0.521*	*0.0014*	*0.522*	*0.615*	*0.727*
30	*5*	*0.0006*	*0.288*	*-0.663*	*-0.559*	*-0.474*	*0.0007*	*0.473*	*0.559*	*0.667*
35	*6*	*-0.0008*	*0.262*	*-0.608*	*-0.514*	*-0.432*	*0.0004*	*0.426*	*0.509*	*0.601*
40	*6*	*0.0012*	*0.243*	*-0.566*	*-0.477*	*-0.400*	*0.0020*	*0.400*	*0.475*	*0.567*
45	*6*	*-0.0015*	*0.235*	*-0.551*	*-0.461*	*-0.386*	*-0.0025*	*0.383*	*0.457*	*0.545*
50	*6*	*0.0002*	*0.219*	*-0.508*	*-0.426*	*-0.361*	*0.0015*	*0.358*	*0.426*	*0.506*
60	*6*	*0.0001*	*0.203*	*-0.469*	*-0.397*	*-0.332*	*-0.0010*	*0.337*	*0.397*	*0.469*
70	*7*	*-0.0002*	*0.181*	*-0.417*	*-0.354*	*-0.298*	*-0.0012*	*0.300*	*0.358*	*0.432*
80	*7*	*-0.0003*	*0.170*	*-0.395*	*-0.330*	*-0.281*	*0.0017*	*0.279*	*0.333*	*0.390*
90	*7*	*0.0002*	*0.160*	*-0.377*	*-0.317*	*-0.265*	*0.0003*	*0.265*	*0.316*	*0.373*
100	*7*	*-0.0004*	*0.152*	*-0.354*	*-0.298*	*-0.250*	*-0.0003*	*0.250*	*0.299*	*0.354*

4 Power of the L.C.S. test against particular alternatives

Supposing that the data we are dealing with are coming from a continuous and unimodal distribution, we can test symmetry by using the Gaussian as the null-hypothesis distribution. We have tested the power of the a test of symmetry based on the statistic *b* (l.c.s.), using Table 2 "modified" tail percentiles as critical values. We considered four positively skewed different alternatives, evaluating their degree of skewness in terms of the classic Pearson parameter of skewness $\sqrt{\beta_1} = \frac{\overline{\mu}_3}{\sigma^3}$ ($\overline{\mu}_3$= third moment about the expected value, σ = std. deviation). We chose these alternative models:

- X_1 : Exponential distribution with parameter λ=1/2, which is the same as a χ^2 with 2 degrees of freedom. This distribution has a value of $\sqrt{\beta_1}$ equal to +2.
- X_2 : Gamma distribution with α=4, β=2. The skewness is $\sqrt{\beta_1}$ = +1.
- X_3 : Log-normal distribution, derived by exponentiating a normal distribution with μ=0 and σ=1/6. The resulting distribution has approximately $\sqrt{\beta_1}$ = +0.5.
- X_4: Skew-normal distribution (see Azzalini, 1985, 1986) with ρ=2/3, with $\sqrt{\beta_1}$ approximately equal to +0,18.

Table 4: Average, standard deviation and power of the l.c.s. statistic for exponential and gamma alternatives.

	Exponential alternative					***Gamma alternative***				
n	*Ave-rage*	*Std. dev.*	*Power* α*=5%*	*Power* α*=2.5%*	*Power* α*=1%*	*Aver-age*	*Std dev.*	*Power* α*=5%*	*Power* α*=2.5%*	*Power* α*=1%*
20	***1.147***	***0.386***	***88.10%***	***81.06%***	***70.75%***	***0.604***	***0.403***	***43.26%***	***31.68%***	***20.87%***
25	***1.281***	***0.388***	***94.06%***	***89.42%***	***81.97%***	***0.674***	***0.402***	***51.46%***	***39.17%***	***27.24%***
30	***1.382***	***0.392***	***96.95%***	***94.05%***	***88.68%***	***0.737***	***0.395***	***59.52%***	***47.71%***	***33.86%***
35	***1.437***	***0.385***	***98.51%***	***96.74%***	***92.62%***	***0.766***	***0.379***	***64.95%***	***52.67%***	***38.76%***
40	***1.510***	***0.390***	***99.22%***	***98.09%***	***95.58%***	***0.812***	***0.378***	***69.79%***	***58.16%***	***44.34%***
45	***1.586***	***0.394***	***99.57%***	***98.99%***	***97.49%***	***0.856***	***0.382***	***74.39%***	***63.53%***	***50.08%***
50	***1.627***	***0.392***	***99.85%***	***99.47%***	***98.57%***	***0.878***	***0.370***	***78.61%***	***68.28%***	***53.45%***
60	***1.731***	***0.401***	***99.95%***	***99.84%***	***99.43%***	***0,941***	***0.372***	***84.40%***	***75.72%***	***62.67%***
70	***1.794***	***0.403***	***99.98%***	***99.93%***	***99.76%***	***0.975***	***0.363***	***88.75%***	***80.47%***	***68.70%***
80	***1.874***	***0.415***	***99.99%***	***99.98%***	***99.94%***	***1.021***	***0.365***	***91.92%***	***85.85%***	***74.60%***
90	***1.951***	***0.421***	***100.00***	***99.98%***	***99.95%***	***1.064***	***0,364***	***94.19%***	***88.85%***	***79.65%***
100	***1.979***	***0.413***	***100.00***	***100.00***	***99.99%***	***1.086***	***0.355***	***96.05%***	***92.11%***	***84.60%***

We evaluated the power corresponding to three different significance levels (α=5%, α=2.5% and α=1%) by performing a series of Monte Carlo simulations. For each alternative distribution and sample size (from 20 to 100, as before), we simulated 20,000 samples and registered the number of them leading to reject the null hypothesis in an α-sized one-tailed test. The frequency of rejection may be regarded as an unbiased estimation of the power of the test. In Table 3 we have represented the average, standard deviation and power of the l.c.s. test for the "highly skewed" alternatives X_1 and X_2, while in Table 4 we reported the same information for the "slightly skewed" alternatives X_3 and X_4.

Looking at Tables 4 and 5, we can notice that the average values of the *l.c.s.* are increasing with *n,* for all the alternatives considered here. The standard deviation shows a slightly increasing trend for an exponential variate, and a slightly decreasing trend for the remaining alternatives. The power of the test, obviously, depends on the sample size and the degree of skewness of the alternative. For Exponential alternative we have a very high power even for small values of *n*; when considering a Gamma alternative we need a larger sample to reach the same power: a sample of size 70 for Gamma alternative is almost equally powerful than a sample of size 20 for Exponential alternative. The same happens with Gamma and Log-normal: we observed almost the same power by considering *n=20* for Gamma and *n=70* for Log-normal. The difference between Log-normal and Skew-normal is just a little bit greater: *n=20* for Log-normal has the same power than *n=75* for Skew-normal.

The power of the test based on the *l.c.s.* has been then compared with the power of the test based on the sample Pearson index $\sqrt{b_1}$. For determining the tail values of $\sqrt{b_1}$ under normality, we made an analogous simulation, generating

30,000 standard normal samples for each value of *n*. We tested our results by comparing them with the tabulated values specified in D'Agostino and Tietjen (1973), for n less or equal to 35: the simulated tail percentiles, reported in Table 6, seem to be consistent with the corresponding tabulated ones.

Table 5: Average, standard deviation and power of the l.c.s. statistic for log-normal and skew-normal alternatives.

	Log-normal alternative					Skew-normal alternative				
n	Ave-age	Std. dev.	Power α=5%	Power α=2.5%	Power α=1%	Aver-age	Std. dev.	Power α=5%	Power α=2.5%	Power α=1%
20	0.298	0.416	18.91%	11.84%	6.28%	0.156	0.407	10.32%	5.73%	2.52%
25	0.342	0.410	21.98%	14.22%	7.86%	0.172	0.398	11.18%	6.00%	2.80%
30	0.369	0.401	25.40%	16.77%	9.75%	0.189	0.387	12.44%	6.96%	3,28%
35	0.386	0.385	28.02%	18.63%	10.74%	0.200	0.372	13.48%	7.67%	3.56%
40	0.410	0.380	29.53%	20.23%	12.30%	0.214	0.364	14.12%	7.98%	3.78%
45	0.433	0.376	32.98%	22.87%	14.00%	0.218	0.359	14.52%	8.64%	3.99%
50	0.444	0.368	35.27%	24.75%	14.75%	0.223	0.347	15.00%	8.89%	4.16%
60	0.473	0.361	38.58%	27.98%	17.62%	0.232	0.344	16.02%	9.48%	4.58%
70	0.492	0.351	42.46%	30.58%	19.77%	0.245	0.330	17.56%	10.53%	5.41%
80	0.520	0.350	46.76%	35.19%	23.03%	0.253	0.325	18.68%	11.34%	5.79%
90	0.545	0.350	49.34%	37.69%	25.93%	0.260	0.324	19.35%	12.11%	6.18%
100	0.554	0.338	52.83%	40.56%	27.52%	0.264	0.309	20.47%	12.44%	6.33%

Table 6: Limit values for the test based on Pearson $\sqrt{b_1}$ statistic.

	Limit values for $\sqrt{b_1}$ (tabulated)			Limit values for $\sqrt{b_1}$ (simulated)		
n	α =5%	α =2.5%	α =1%	α =5%	α=2.5%	α =1%
20	0.777	0.951	1.152	0.776	0.943	1.152
25	0.714	0.876	1.073	0.710	0.869	1.060
30	0.664	0.804	0.985	0.667	0.810	1.003
35	0.624	0.762	0.932	0.630	0.760	0.941
40	- - -	- - -	- - -	0.584	0.712	0.868
45	- - -	- - -	- - -	0.558	0.679	0.822
50	- - -	- - -	- - -	0.529	0.640	0.780
60	- - -	- - -	- - -	0.491	0.595	0.726
70	- - -	- - -	- - -	0.456	0.547	0.658
80	- - -	- - -	- - -	0.428	0.521	0.628
90	- - -	- - -	- - -	0.405	0.488	0.586
100	- - -	- - -	- - -	0.389	0.467	0.567

We have then calculated the power of the test based on $\sqrt{b_1}$ by executing an analogous simulation of 20,000 samples the same size and compared, for each alternative, the power of the l.c.s. and $\sqrt{b_1}$ statistics. In Figures 2, 3, 4, 5 we give a graphical representation of such a comparison: we have plotted the ratio $R(n.\alpha) = \dfrac{PL\,(n,\alpha)}{PM\;(n,\alpha)}$, where *PL(n,α)* is the power of the l.c.s. test (L stands for Letter values), referred to a sample size n and a significance level α, and *PM(n,α)* is the corresponding the power of the classic $\sqrt{b_1}$ test (M stands for Moments). When *R(n,α)* is greater than one, the l.c.s. test is more powerful; when it is less than one, the new test is less powerful.

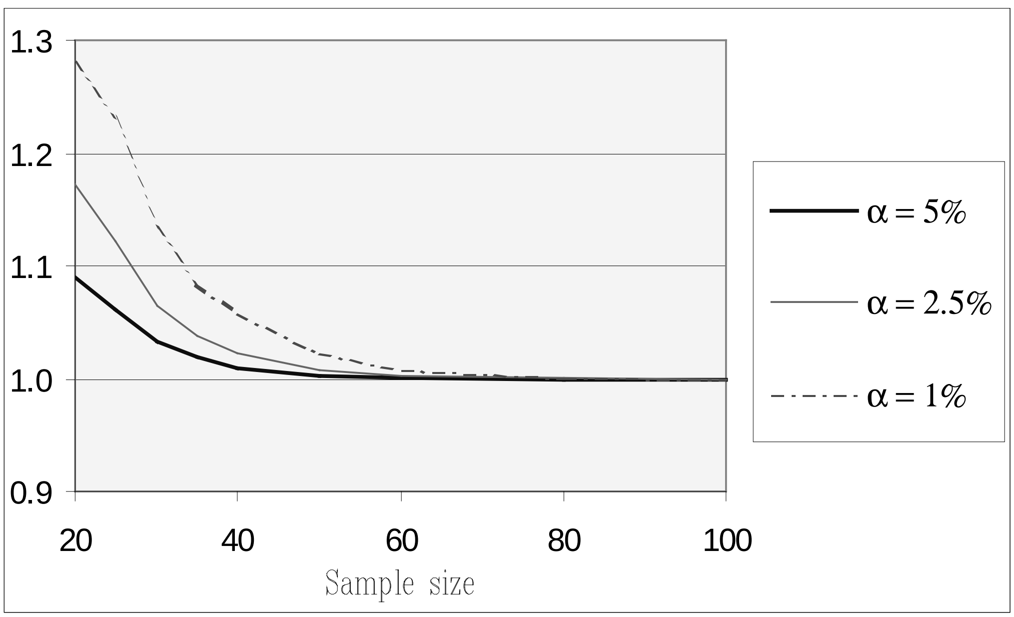

Figure 2: Ratio between the power of the l.c.s. test and the standard $\sqrt{b_1}$ test against *an exponential alternative.*

The simulation results show that the l.c.s. test is more powerful for such a strongly skewed alternative, for each value of α and n considered. This is due to the particular stress that l.c.s. gives to the tail behavior. The difference between the statistics is greater for small values of n; when $n \geq 60$, the two tests have virtually the same power.

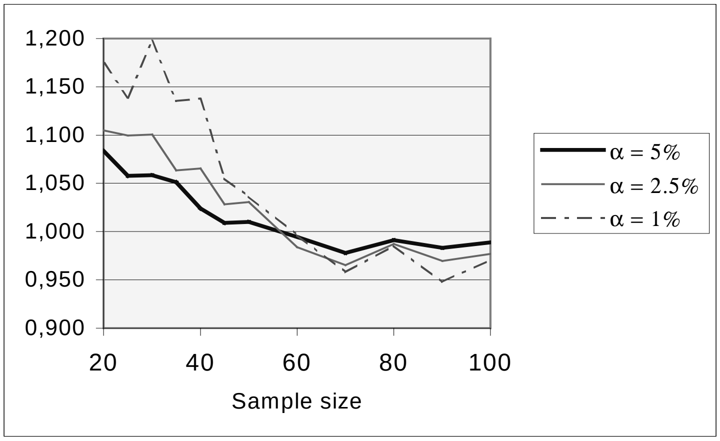

Figure 3: Ratio between the power of the l.c.s. test and the standard $\sqrt{b_1}$ test against a Gamma (4,2) alternative.

Looking at the graph, we notice for this alternative that the l.c.s. test is more powerful (especially when the significance level α is small), for a reduced sample size (from 20 to 50). When n is equal to greater than 60, there is not a marked difference in power; however, the test based on $\sqrt{b_1}$ is slightly more powerful.

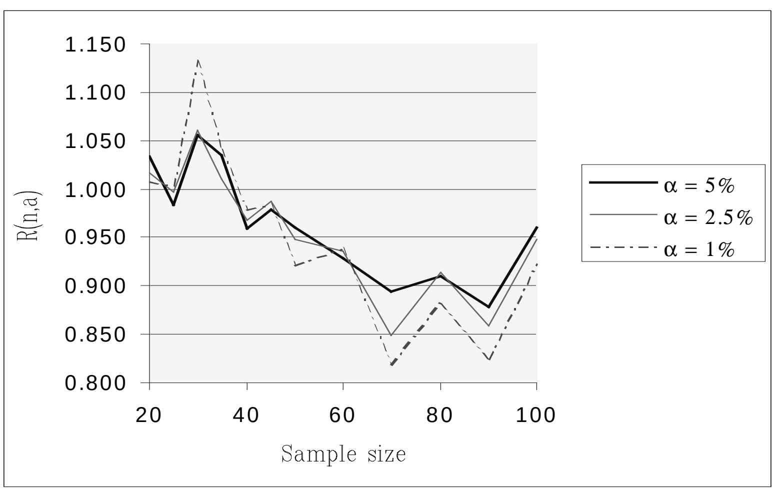

Figure 4: Ratio between the power of the l.c.s. test and the standard $\sqrt{b_1}$ test against a Log-normal alternative.

Here, with a not-too-skewed Log-normal alternative, the two tests have almost the same power for n=20, 25, while the l.c.s. test results to be more powerful for n = 30, 35. When the sample size is larger, the classic test based on moments is sensitively more powerful than the l.c.s. test.

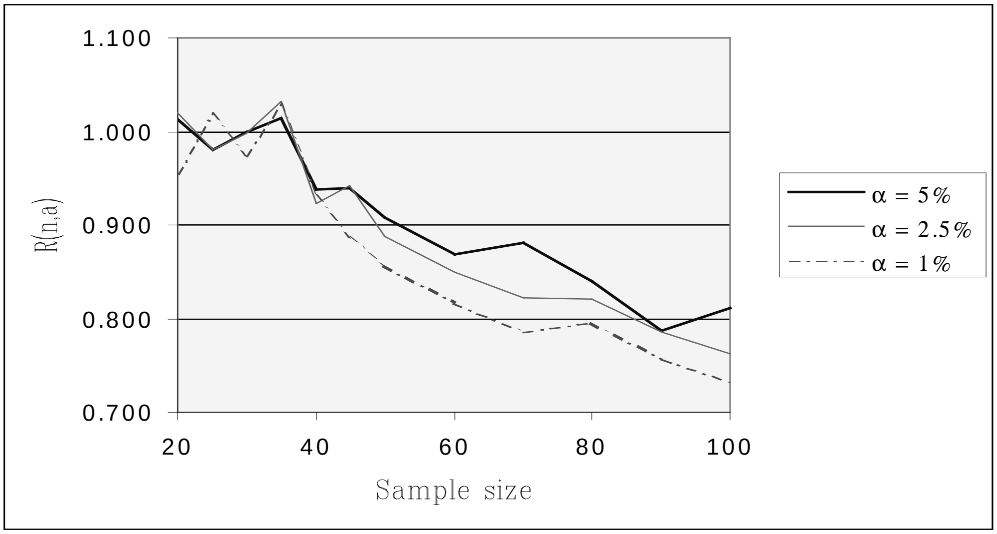

Figure 5: Ratio between the power of the l.c.s. test and the standard $\sqrt{b_1}$ test against a Skew-normal alternative.

With a slightly-skewed skew normal alternative, the l.c.s test has almost the same power as the classic test, but the ratio R(n,α) becomes very low when the sample size increases.

Then, the l.c.s. works well when the alternative is strongly skewed, especially when the sample is not too large and for small values of α. Therefore, the l.c.s. test seems to be suitable as a quick test for detecting if the sample comes from a markedly skewed population.

5 Concluding remarks

The *l.c.s.*, as pointed out in the title, has the advantage to be easy (and quick) to calculate: it needs only to detect a reduced number of order statistics, to standardize such values and to make a least squares linear interpolation. As described in Section 4), the test based on this index of skewness, being focused on sample tails, seems to be more powerful than the standard test based on Pearson statistic $\sqrt{b_1}$, whenever the sample comes from a population with a certain degree of skewness, and particularly when the sample is not too large (from 20 to 50). Moreover, the *l.c.s.* is invariant under linear transformations of the data, with the only condition that the slope is positive. This allows to make comparisons between different sets of data, regardless of the scale of measurement.

There are still many things to do on this topic. The next step of this research could be the definition of a robust version of the index, by trimming the sample data or the set of letter values, in order to reduce the influence of the extremes. In Brizzi (2000) there is a proposal in this sense. The power of the test, standard and robustified, may be compared with other tests of symmetry, e.g., the tests described in Antille et al. (1982); moreover, it would be important to study the sample distribution, finite and asymptotic, by an analytical point of view, in order to support the simulated results.

References

[1] Antille, A., Kersting, G., and Zucchini, W. (1982): Testing symmetry. *Journal of the American Statistical Association*, **77**, 639-646.

[2] Azzalini, A. (1985): A class of distributions which includes the normal ones. *Scandinavian Journal of Statistics*, **12**, 171-178.

[3] Azzalini, A. (1986): Further results on a class of distributions which includes the normal ones. *Statistica*, **XLVI**, 199-208.

[4] Balanda, K.P. and Mac Gillivray, H.L. (1988): Kurtosis: a critical review. *The American Statistician*, **42**, 111-119.

[5] Brizzi, M. (2000): Detecting skewness and kurtosis by letter values: a new proposal. *Statistica***, LX,** 243-258.

[6] D'Agostino, R.B. and Tietjen, G.L. (1973): Approaches to the null distribution of $\sqrt{b_1}$. *Biometrika*, **60**, 169-173.

[7] Groeneveld, R.A. (1998): A class of quantile measures for kurtosis. *The American Statistician*, **52**, 325-329.

[8] Hoaglin, D.C., Mosteller, F., and Tukey, J.W. (1983): *Understanding Robust and Exploratory Data Analysis*. New York: John Wiley & Sons.

[9] Hoaglin, D.C., Mosteller, F., and Tukey, J.W. (1985): *Exploring Data Tables, Trends and Shapes*. New York: John Wiley & Sons. Chapter 10 by D.C. Hoaglin.

[10] Kappenman, R.F. (1988): Detection of symmetry or lack of it and application. *Communications in Statistics A. Theory and Methods*, **17**, 4163-4177.

[11] Mac Gillivray, H.L. (1986): Skewness and asymmetry: measures and orderings. *Annals of Statistics*, **14**, 994-1011.

[12] Moors, I.J.A. (1988): A quantile alternative for kurtosis. *The Statistician*, **37**, 25-32.

[13] Oja, H. (1981): On location, scale, skewness and kurtosis. *Scandinavian Journal of statistics*, **8**, 154-168.

[14] Ruppert, D. (1987): What is Kurtosis? *The American statistician*, **41**, 1-5.

[15] Tukey, J.W. (1977): *Exploratory Data Analysis*. Reading: Addison - Wesley.

[16] Zenga, M. (1996): La curtosi. *Statistica*, **LVI**, 87-102.