International Journal of Business Management and Economics Research.
ISSN 2349-2333 Volume 2, Number 2 (2015), pp. 95-106

http://www.irphouse.com

Parametric Test for Skewness from Unknown Distributions

P.V.Pushpaja

Nehru Arts & Science College
Kanhangad-671328
Kerala, India

Abstract

Here we present a new parametric method of testing skewness using order statistics. By deriving the first four central moments of the test statistics, the distribution of it is fitted as a member in the Generalized Lambda Distribution (GLD) family. The method is based on computer programmes in Maple language. Tables representing the percentile points of the sampling distribution of the test statistics under different distributions and for different sample sizes are provided. The power of the test is also obtained. This method is too general as there is no specific distributional assumption.

Key words: Order statistics, GLD family, Maple programmes, Parametric tests, Moments, Skewness and Kurtosis measures.

1. Introduction

A random variable X with probability density function (pdf) f(x) is said to be symmetric about a point ‘a’, if f(a + x) = f(a − x). A distribution that lacks symmetry is said to be skewed. Some of the commonly used measures of skewness are listed below.

(1). Karl Pearson measure of skewness

ξ=(mean – mode)/standard deviation

ξ_1=(mean – median)/standard deviation, when the mode is not easily available.

If μ_r, r = 2 to 4 denote the central moments of order 2 to 4 of X, then the measure of skewness based on moments is $\beta_1 = \mu_3^2 / \mu_2^3$

If Q1, Q2 and Q3 denote the first three quartiles of the distribution of a random variable X, then the quantile measure of skewness is denoted as β and is defined as

β = (Q3 − Q2) − (Q2 − Q1).

If m_r, r = 2, 3 & 4 denote the sample central moments of order 2, 3 & 4, then the sample measure of skewness based on moments is $b_1 = m_3^2 / m_2^3$ and the kurtosis is $b_2 = m_4 / m_2^2$.

Several attempts have been made to fit a distribution to m3 or to the standardized moment ratio b_1, based on sample observations from a normal population. Using the observations from a normal population, Pepper (1932) fitted a Pearson type VII distribution to the distribution of m_3. While testing goodness of fit, Pepper found that the addition of one value of m3 at a sufficient distance from the origin would greatly increase the value of the standardized moment ratios $\beta_3 = \mu_6 / \mu_2^3$ and $\beta_4 = \mu_8 / \mu_2^{4.}$ When sampling is from a normal population, Pearson (1965) suggested a Johnson's symmetric Su curve for the distribution of b1 and D'Agostino and Pearson (1973) provided tables giving the values of probability integrals. A test based on a set of observations to see whether they could have arisen by random sampling from a normal population is termed as a test of normality, which may be carried out by a comparison of the sample distribution function with a normal distribution function. Certain other tests are said to be tests of normality when they are only tests of agreement of certain sample statistics with the values of the corresponding normal parameters. For example, a test of the value of the moment ratio b_1 against the normal value of zero or a test of the value of b_2 against the normal value 3. Based on a test-statistic W, which is the ratio of the square of the best linear unbiased estimator of the normal population standard deviation to the sample variance, Shapiro and Wilk (1965) suggested a test of normality which approximately detect whether the deviation from normality is due to skewness or kurtosis. They supplied tables representing the percentile points of the null distribution of W for sample sizes 3 to 50. The above procedures can be applied if the observed data is from a normal population and for large samples only. Even when the normality assumption holds good for observations, it is difficult to find the distribution of the test statistics and hence power of the test cannot be determined.

In this article, we suggest a general parametric test for the quantile measure of skewness. Based on the first four moments of the test statistics, its sampling distribution is fitted as a member in the generalized lambda distribution (GLD) family. The critical region and power of the test are also evaluated. To find the first four moments of the test statistics, to fit its sampling distribution and to find power of the test, computer programmes in Maple language are provided. The method can be applied in the case of small samples also. Since there is no assumption regarding the form of the parent distribution except continuity of the random variables, it can be used for all type of distributions (symmetric, positively skewed and negatively skewed). The most significant aspect of the method is that, by inputting the given set of observations in the programme, one can examine whether the observations are drawn from a population with the specified skewness measure against any of the alternatives. Also, the power values under any of the alternatives can be evaluated. A brief review of the GLD family is discussed in section 2. In section 3, the first four moments of the test statistics are derived and the test is described. Illustration of the method on different type of distributions together with tables representing the percentile points of the distribution of the test statistics and power values for different sample sizes are provided in section 4. Real life examples are also provided in this section. The programmes to find the first four moments of the test statistics, to fit its distribution, to find critical region and power of the test are given in appendix.

2. Generalized Lambda Distributions (GLD) Family

The generalized lambda distribution (GLD) family is a four parameter family of distributions derived by Ramberg and Schmeiser (1974). Unlike most other four parameter family of distributions, GLD has no explicit expression for its pdf, instead, members of the family are specified in terms of their quantile function. The quantile function of the four parameter GLD family is given by

$Q(p) = \lambda_1 + [p^{\lambda 3} - (1\text{-}p)^{\lambda 4}]/ \lambda_2$

Here, λ_1 and λ_2 represent the location and scale parameters where as λ_3 and λ_4 represent the shape parameters of the distribution. The support of the random variable with the above distribution is $[\lambda_1 - 1/\lambda_2, \lambda_1+1/\lambda_2]$ when $\lambda_3>0$ and $\lambda_4>0$. The support is $(-\infty, \lambda_1+1/\lambda_2)$ when $\lambda_3 < 0$ and $\lambda_4 = 0$ and it is $(\lambda_1 - 1/\lambda_2, \infty)$ when $\lambda_3 = 0$ and $\lambda_4 < 0$.

2.1 Estimation of Parameters and Fitting of GLD

The popular method of fitting GLD to a data set is the method of moments due to Ramberg et al. (1979). In this method the parameters λ_3 and λ_4 are first derived by solving the equations $\alpha_3 = a_3$ and $\alpha_4 = a_4$, where α_3 and α_4 are the coefficients of skewness and kurtosis of the distribution and a_3 and a_4 are their sample estimates. These systems of equations are too complex, so that to obtain the solutions computer programme in Maple language is provided. By solving the equations $\alpha_1 = a_1$, $\alpha_2 = a_2$ and using the estimated values of λ_3 and λ_4, the values of λ_1 and λ_2 were determined. It may be noted that skewness and kurtosis are independent of location and scale parameters and moments of all orders exist if λ_3 and λ_4 are of same sign. Since corresponding to every admissible pair of skewness and kurtosis measures GLD family contains a member, a wide variety of densities with different tail shapes are available in the family. This family was used for Monte-Carlo simulation studies of robustness of statistical procedures and for sensitivity analysis. The family contains unimodal, U-shaped, J-shaped, symmetric and asymmetric distributions. One of the important advantages of this family is that all its members can be represented by a single quantile function and almost all known distributions can be represented as its member.

3 Test for Skewness

Let X be a unimodal continuous random variable with distribution function F(x), then the p^{th} quantile of X is a real valued function denoted as Q(p) and is defined as

$$Q(p) = \inf(t : F(t) \geq p),\ 0 < p < 1, \qquad (3.1)$$

Then the first three quartiles Q1, Q2 and Q3 are obtained by putting p as ¼, ½ and ¾ respectively in equation (3.1).

Let $x_1, x_2,... x_n$ denote a random sample of size n drawn from a population with the quantile measure of skewness as β and let $x_{1;n}, x_{2;n},... x_{n;n}$ denote the order statistics.

Then for every p, $0 < p < 1$, the p^{th} quantile in the sample is denoted as q(p) and is defined as

$q(p) = x_{np:n}$, if np is integer

$= x_{[np]+1:n}$, otherwise, where [np] is the integer part of np. (3.2)

Let q1, q2 and q3 denote the first three sample quartiles, which are obtained by putting p as ¼, ½ and ¾ respectively in equation (3.2), then the sample estimate of β is

$b = (q3 - q2) - (q2 - q1)$ (3.3)

To test H0 : β =β0 against any of the alternatives H1 : β >β0, H1 :β < β0 or H1 : β≠ β0, the test statistics used is b (β0 is a specified value of β).

Detailed description regarding the theory and development of order statistics are available in David (1981) and David and Nagaraja (2003). Balakrishnan and Rao (1998) and Harter and Balakrishnan (1996) described the recent developments in the theory and properties of order statistics. Reiss (1989) described in details the relation between the quantiles and order statistics.

3.1 First Four Moments of the Test-Statistics

Let $\mu_r^{'}(k:n)$, r = 1, 2,... denote the moments about origin of the k^{th} order statistics xk:n, based on a random sample of size n drawn from the above population. Kumaran and Beena (2005) derived general expression for moments about origin of order statistics from GLD family,

$\mu_r^{'}(k:n) = (1/\beta(k, n-k+1)) \sum_{i=0}^{r} \binom{r}{i} \lambda_1^{r-I} \lambda_2^{-I} \sum_{j=0}^{i} (-1)^j \binom{i}{j} \beta(\lambda_3 (i-j)+k, \lambda_4 j + n - k + 1)$ (3.4)

Putting r = 1, 2, 3, 4 in equation (3.4) and simplifying we obtain the first four raw moments of k^{th} order statistics as

$\mu_1^{'}(k:n) = \lambda_1 + [A_1 / (A_0\lambda_2)]$ (3.5)

$\mu_2^{'}(k:n) = \lambda_1^2 + [(2\lambda_1 A_1)/(A_0\lambda_2)] + [A_2 / (A_0\lambda_2^2)]$ (3.6)

$\mu_3^{'}(k:n) = \lambda_1^3 + [(3\lambda_1^2 A_1)/(A_0\lambda_2)] + [3 A_2 \lambda_1 / (A_0\lambda_2^2)] + [A_3 / (A_0\lambda_2^3)]$ (3.7)

$\mu_4^{'}(k:n) = \lambda_1^4 + [(4\lambda_1^3 A_1)/(A_0\lambda_2)] + [6 A_2 \lambda_1^2/(A_0\lambda_2^2)] + [4\lambda_1 A_3 / (A_0\lambda_2^3)] + [A_4 / (A_0\lambda_2^4)]$ (3.8)

where $A_0 = \beta(k, n-k+1)$ (3.9)

$A_1 = \beta(\lambda_3 + k, n-k+1) - \beta(k, \lambda_4 + n - k + 1)$ (3.10)

$A_2 = \beta(2\lambda_3 + k, n-k+1) - 2\beta(\lambda_3 + k, \lambda_4 + n - k + 1) + \beta(k, 2\lambda_4 + n - k + 1)$ (3.11)

$A_3 = \beta(3\lambda_3 + k, n-k+1) - 3\beta(2\lambda_3 + k, \lambda_4 + n - k + 1) + 3\beta(\lambda_3 + k, 2\lambda_4 + n - k + 1) - \beta(k, 3\lambda_4 + n - k + 1)$ (3.12)

$A_4 = \beta(4\lambda_3 + k, n-k+1) - 4\beta(3\lambda_3 + k, \lambda_4 + n - k + 1) + 6\beta(2\lambda_3 + k, 2\lambda_4 + n - k + 1) - 4\beta(\lambda_3 + k, 3\lambda_4 + n - k + 1) + \beta(k, 4\lambda_4 + n - k + 1)$ (3.13)

Hence, the central moments up to order four of k^{th} order statistics are obtained as

$\mu_2(k:n) = \lambda_2^{-2} [(A_2/A_0) - (A_1^2/A_0^2)]$ (3.14)

$\mu_3(k:n) = \lambda_2^{-3} [(A_3/A_0) - (3A_1 A_2/A_0^2) + (3A_1^3/A_0^3)]$ (3.15)

$\mu_4(k:n) = \lambda_2^{-4} [(A_4/A_0) - (4A_1 A_3/A_0^2) + (6A_1^2 A_2/A_0^3) - (3A_1^4/A_0^4)]$ (3.16)

Since the distribution function of sample observations is the estimator of the population distribution function, the sample quantile function can be regarded as an estimator of the population quantile function. Hence the mean, variance, skewness and kurtosis of the test statistics b are respectively denoted as α_1, α_2, α_3, α_4 and are obtained as

$$\alpha_1=\beta \tag{3.17}$$

$$\alpha_2 = \mu_2(q_3) + 4\mu_2(q_2) + \mu_2(q_1) \tag{3.18}$$

$$\mu_3(b) = \mu_3(q_3) - 8\mu_3(q_2) + \mu_3(q_1) \tag{3.19}$$

$$\mu_4(b) = \mu_4(q_3) + 10\mu_4(q_2) + \mu_4(q_1) + 24\mu_2(q_3)\mu_2(q_2) + 6\mu_2(q_3)\mu_2(q_1) + 24\mu_2(q_2)\mu_2(q_1) + 6[\mu_2(q_2)]^2 \tag{3.20}$$

$$\alpha_3 = \mu_3(b)/(\mu_2(b))^{3/2}$$

$$\alpha_4 = \mu_4(b)/(\mu_2(b))^2$$

The values of $\mu_2(q_1)$, $\mu_3(q_1)$ and $\mu_4(q_1)$ are respectively obtained by putting k as n/4 or [n/4] +1 according as n/4 is an integer or not in equations (3.14) to (3.16), where [n/4] represents the integer part of n/4. Similarly, the moments of q2 and q3 are obtained by putting the value of k as n/2 or [n/2] +1 and 3n/4 or [3n/4] +1 as the case may be in equations (3.14) to (3.16).

3.2 The Test for Skewness

To test (H0 :$\beta = \beta_0$) against any of the alternatives, the test procedure is described below in the form of an algorithm.

Step-1. Computation of test statistics: Draw a sample of size n from the given population with GLD parameters (λ_1, λ_2, λ_3, λ_4). Arranging the sample observations in their order of magnitude, compute the three sample quartiles q_k, k = 1, 2, 3, which are the $(nk/4)^{th}$ observations in the ordered sample when nk/4 is an integer and are $([nk/4] + 1)^{th}$ observations when nk/4 is not an integer. Then compute the test statistics $b = (q_3 - q_2) - (q_2 - q_1)$.

Step-2. Computation of first four moments of the test statistics: The mean, variance, skewness and kurtosis of the test statistics are obtained using equations (3.14) to (3.20). To obtain these values, programme P1 in the appendix can be used and denote them respectively as α_1, α_2, α_3, α_4.

Step-3. Fitting GLD to the distribution of the test statistics: Using these values of α_1, α_2, α_3, α_4 as the arguments in the programme P2 in the appendix, fit the distribution of the test statistics under H0 as a member in the GLD family. Let (λ_1', λ_2', λ_3', λ_4') denote the GLD parameters of the fitted distribution and hence the quantile function of the fitted distribution of b under H_0 is given as

$$Q(p) = \lambda_1' + [p^{\lambda_3'} - (1-p)^{\lambda_4'}]/ \lambda_2', 0\leq p \leq 1 \tag{3.21}$$

Step-4. Computation of critical region: For a given level of significance α, the critical regions are $b \leq S1$, for $H1 : \beta < \beta_0$, $b \geq S2$, for $H1 : \beta > \beta_0$ and $b \leq S3$ or $b \geq S4$, for $H1 : \beta \neq \beta_0$, where S1, S2, S3 and S4 are obtained from equation (3.21) by putting p as α, $1-\alpha$, $\alpha/2$ and $1-\alpha/2$ respectively.

Step-5. Computation of power: To obtain the power of the test, for different values of β, re-estimate the GLD parameters of the distribution of b and from the fitted distribution the required probabilities can be evaluated using the programme P3.

Remark:-In case when the GLD parameters of the parent distribution are unknown, they can be estimated using the sample observations and programme P2.

4 Numerical Illustration

The method proposed above for testing the significance of skewness is illustrated in the case of log-normal (0, 1/3) distribution, which is a positively skewed distribution with the GLD parameters $\lambda_1 = 0.8451$, $\lambda_2 = 0.1085$, $\lambda_3 = 0.01017$ and $\lambda_4 = 0.03422$. To test H0 : $\beta = 0.03875$, the values of (α_1, α_2, α_3, α_4) and the corresponding GLD parameters of the distribution of the test statistics for different values of n are presented in table 1. For different values of n, the percentile points of the distribution of b and the power values at $\alpha = 0.05$ are presented in tables 2 and 3. Based on these percentile points, conclusions can be made against any of the alternatives, H1 : $\beta > 0.03875$, H1 : $\beta < 0.03875$ or H1 : $\beta \neq 0.03875$

Table 1: The values of (α1, α2, α3, α4) and the GLD parameters of the distribution of b from LN(0, 1/3)

n	α_1	α_2	α_3	α_4	λ_1	λ_2	λ_3	λ_4
5	0.0387	0.2062	-0.2191	3.4058	-0.0381	0.2728	0.0656	0.0899
10	0.0387	0.1031	-0.0784	3.2194	0.0162	0.4767	0.0927	0.1057
15	0.0387	0.0740	-0.0711	3.1822	0.0208	0.5890	0.0982	0.1111
20	0.0387	0.0514	-0.0781	3.0738	0.0202	0.8033	0.1128	0.1316
25	0.0387	0.0430	-0.0719	3.1624	0.0246	0.7911	0.1008	0.1146
30	0.0387	0.0356	-0.0525	3.0519	0.0281	0.9881	0.1190	0.1324

Table 2: Percentile points of the distribution of b from LN(0, 1/3)

	% Probability Level							
n	90	95	97.5	99	10	5	2.5	1
5	0.6221	0.8150	0.9904	1.2021	-0.5171	-0.6748	-0.8171	-0.9899
10	0.4490	0.5756	0.6886	0.8227	-0.3638	-0.4812	-0.5858	-0.7105
15	0.3866	0.4930	0.5875	0.6992	-0.3031	-0.4025	-0.4904	-0.5950
20	0.3308	0.4184	0.4952	0.5844	-0.2473	-0.3283	-0.3993	-0.4824
25	0.3044	0.3854	0.4572	0.5417	-0.2222	-0.2976	-0.3644	-0.4435
30	0.3044	0.3854	0.4572	0.5417	-0.2222	-0.2976	-0.3644	-0.4435

Table 3: Power values of the distribution of the test from LN(0, 1/3)

N	β: for lower tailed test					β: for upper tailed test				
	-0.2	-0.4	-0.6	-0.8	-1	0.2	0 .4	0.6	0.8	1
5	0.088	0.170	0.298	0.468	0.652	0.138	0.272	0.448	0.630	0.778
10	0.118	0.282	0.524	0.760	0.910	0.186	0.404	0.654	0.846	0.946
15	0.136	0.356	0.652	0.874	0.970	0.226	0.504	0.774	0.928	0.984
20	0.164	0.460	0.790	0.956	0.9952	0.288	0.632	0.886	0.980	0.998
25	0.180	0.522	0.854	0.976	0.998	0.320	0.698	0.928	0.990	0.999
30	0.204	0.606	0.906	0.991	1	0.360	0.762	0.960	0.998	1

4.1 Application of the Method

The test developed in this section is illustrated on real life data given below.

Example-1: The following data represents the scores on intelligence quotient (I.Q) examinations of 40 sixth grade students at a particular school.

114, 122, 103, 118, 99, 105, 134, 125, 117, 106, 109, 104, 111, 127, 113, 121, 133, 111, 117, 103, 120, 98, 100, 130, 141, 119, 128, 106, 109, 115, 100, 130, 125, 117, 119, 113, 104, 108, 110, 102.

Examine whether the I.Q distribution is symmetric or not(Ross (2005)).

Solution: The hypothesis to be tested in this case is H0 : $\beta = 0$ against H1 : $\beta \neq 0$. To conduct the test, given sample is ordered as

98, 99, 100, 100, 102, 103, 103, 104, 104, 105, 106, 106, 108, 109, 109, 110, 111, 111, 113, 113, 114, 115, 117, 117, 117, 118, 119, 119, 120, 121, 122, 125, 125, 127, 128, 130, 130, 133, 134, 141.

Here, the first three quartiles are respectively obtained as q1 = 105, q2 = 113 and q3 = 121. Hence, the observed value of b is obtained as 0. The values of $\alpha_1, \alpha_2, \alpha_3, \alpha_4$ and the corresponding GLD parameters of the distribution of b under H0 are obtained as $\alpha_1 = 0$, $\alpha_2 = 0.1681$, $\alpha_3 = -0.0016$, $\alpha_4 = 2.6251$ and $\lambda_1' = -0.0042$, $\lambda_2' = 0.7028$, $\lambda_3' = 0.2218$, $\lambda_4' = 0.2236$. Since the test is two tailed, at $\alpha = 0.05$, the lower and upper percentile values corresponding to 0.025 are respectively obtained as S3 = −0.7889 and S4 = 0.7897 and the critical region of the test is b < S3 or b > S4. Since the observed value of b falls in the acceptance region, the null hypothesis of symmetry is accepted. Thus the given data support the assertion that the I. Q. distribution is symmetric.

Example-2: Dudewicz, et al. (Karian and Dudewicz(2000)) gave data on the brain tissue MRI scan parameter, AD. It should be noted that the term ’parameter’ is used differently in brain scan studies. Report on the following 23 observations associated with scan of the left thalamus were obtained.

108.7, 107, 110.3, 110, 113.6, 99.2, 109.8, 104.5, 108.1, 107.2, 112, 115.5, 108.4, 107.4, 113.4, 101.2, 98.4, 100.9, 100, 107.1, 108.7, 102.5, 103.3.

Examine whether the distribution of AD is negatively skewed or not.

Solution: The hypothesis to be tested in this case is H0 : $\beta = 0$ against H1 : $\beta < 0$. To conduct the test, since the form of the distribution is not known, the GLD parameters of the parent distribution are estimated using the sample moments. The estimates of the first four moments are 106.8, 22.3, −0.1615, 2.106 and the corresponding GLD parameters are estimated as λ_1 = 101.52, λ_2 = 0.0635, λ_3 = 0.0535 and λ_4 = 0.6316. Using these values, the values of $\alpha_1, \alpha_2, \alpha_3, \alpha_4$ and the corresponding GLD parameters of the test statistics b under H0 are obtained as α_1 = 0, α_2 = 12.61, α_3 = −0.066, α4 = 2.61 and λ_1' = −0.6367, λ_2' = 0.0812, λ_3' = 0.186, λ_4'= 0.264. Since the test is lower tailed, at α = 0.05, the lower percentile point of the distribution of b is S1 = −5.93. The ordered sample is obtained as

98.4, 99.2, 100, 100.9, 101.2, 102.5, 103.3, 104.5, 107, 107.1, 107.2, 107.4, 108.1, 108.4, 108.7, 108.7, 109.8, 110, 110.3, 112, 113.4, 113.6, 115.5.

The first three quartiles are obtained as q1 = 102.5, q2 = 107.4 and q3 = 110 and hence the value of the test statistics is obtained as b = −2.3. The critical region of the test is b > −5.73. Since the observed value of b falls in the rejection region, the null hypothesis is rejected. Thus the given data do not support the assertion that the AD distribution is negatively skewed. Here, for β = 1, the power of the test is obtained as 0.102.

Conclusion

We have demonstrated a general parametric method of test of skewness, based on samples from any continuous unimodal populations. Since, there doesn't exist any exact method for testing the skewness or the third central moment of a population, this method is more useful in practical cases. Again, since it is possible to fit the distribution of the test statistics under both the null and alternative hypotheses, power of the test can also be evaluated under this method. As this method doesn't make any rigorous assumption on the distribution of the population, is applicable to almost all continuous distributions. This method is can be conducted based on computer programmes in maple language.

Appendix

Programme-P1

Procedure to estimate the values of ($\alpha_1, \alpha_2, \alpha_3, \alpha_4$) of b via observations
Function: Findalphas of the sample median via observations
Purpose:– Compute a-values of the sample quantile measure of skewness
Arguments: a, b, c, d–The GLD parameters of the parent distribution;
p, p1, p2–The proportions to find quartiles;
n–number of observations
H–The specified value of the quantile measure of skewness under the null hypothesis;

Findalphas:= Proc(n::Numeric, p::Numeric, p_1::Numeric, p_2::Numeric a::Numeric, b::Numeric, c::Numeric,d::Numeric, H:Numeric)

Local f, f_1, f_2, k, s, u, B_1, B_2, B_3, B_4, B_5, B_6, B_7, B_8, B_9, B_{10}, B_{11}, B_{12}, B_{13}, B_{14}, B_{15}, C_1, C_2, C_3, C_4, C_5, C_6, C_7, C_8, C_9, C_{10}, C_{11}, C_{12}, C_{13}, C_{14}, C_{15}, D_1, D_2, D_3, D_4, D_5, D_6, D_7, D_8, D_9, D_{10}, D_{11}, D_{12}, D_{13}, D_{14}, D_{15}, M1,1, M2,1, M3,1,M4,1, M2,2, M2,3, M3,2, M3,3, M4,2, M4,3, a_1, a_2, a_3, a_4, Ah;

f := convert((n+1)p, fraction); k := trunc(f); f_1 := convert((n+1)p_1, fraction); s := trunc(f); f_2 := convert((n+1)p_2, fraction); u := trunc(f);
B_1 := evalf(Beta(k, n − k + 1)); C_1 := evalf(Beta(s, n −s + 1)); D_1 := evalf(Beta(s, n − s + 1));
B_2 := evalf(Beta(c+k, n− k+1)); C_2 := evalf(Beta(c+ s, n−s + 1)); D_2 := evalf(Beta(c +u, n −u + 1));
B_3 := evalf(Beta(k, d+ n− k +1)); C_3 := evalf(Beta(s, d + n−s +1)); D_3 := evalf(Beta(u, d + n−u + 1));
B_4:=evalf(Beta(2*c + k, n− k+1)); C_4:=evalf(Beta(2*c+s, n−s+1)); D_4:=evalf(Beta(2*c+u, n−u + 1));
B_5:=evalf(Beta(c+k, d+n−k+1)); C_5:=evalf(Beta(c+s, d+n−s+1)); D_5:=evalf(Beta(c+u, d+n−u+1));
B_6:=evalf(Beta(k,(2*d+n−k+1)));C_6:=evalf(Beta(s, (2*d+n−s+1))); D_6:=evalf(Beta(u, (2*d+n−u+1)));
B_7:=evalf(Beta(3*c+k, n− k+1)); C_7:=evalf(Beta(3*c+s, n−s+1)); D_7:=evalf(Beta(3*c+u, n− u+1)); B_8:=evalf(Beta(2*c+k, d+ n− k+1)); C_8:=evalf(Beta(2*c+s, d+ n− s+1)); D_8:=evalf(Beta(2*c+u, d+ n−u+ 1));
B_9:=evalf(Beta(c+ k, 2d+ n−k+1)); C_9:=evalf(Beta(c+ s, 2d+ n−s+1)); D_9:=evalf(Beta(c+ u, 2d+ n−u+1));
B_{10}:=evalf(Beta(k, 3*d+ n−k+1)); C_{10}:=evalf(Beta(s, 3*d+ n−s+1)); D_{10}:=evalf(Beta(u, 3*d+ n−u+1));
B_{11}:=evalf(Beta(4*c+ k, n− k+1)); C_{11}:=evalf(Beta(4*c+ s, n−s+1)); D_{11}:=evalf(Beta(4*c+ u, n− u+1));
B_{12}:=evalf(Beta(3*c+ k, d+ n− k+1)); C_{12}:=evalf(Beta(3*c+ s, d+ n− s+1)); D_{12}:=evalf(Beta(3*c+u, d+ n−u+1));
B_{13}:=evalf(Beta(2*c+k, 2d+n−k+1)); C_{13}:=evalf(Beta(2*c+s, 2d+n−s+1)); D_{13}:=evalf(Beta(2*c+u, 2d+n−u+1));
B_{14} := evalf(Beta(c+k, 3*d+n−k+1)); C_{14} := evalf(Beta(c+s, 3*d+n−s+1)); D_{14} := evalf(Beta(c+u, 3*d+n−u+1));
B_{15} := evalf(Beta(k, 4*d + n − k + 1)); C_{15} := evalf(Beta(s, 4*d + n −s + 1)); D_{15} := evalf(Beta(u, 4*d + n − u + 1));
$M_{1,1}$:= evalf(a + (B_2−B_3)/bB$_1$); $M_{2,1}$:= evalf(b^{-2} [(B_4−2B_5+B_6)/B_1− (B_2−B_3)^2/B_1^2]);
$M_{3,1}$:= evalf(b^{-3} [(B_7−3B_8 +3B_9-B_{10})/B_1− 3(B_2-B_3)(B_4−2B_5+B_6)/B_1^2 +2 (B_2−B_3)^3/B_1^3]);
$M_{4,1}$:= evalf(b^4 [(B_{11}−4B_{12} +6B_{13}-4B_{14} +B_{15})/B_1− 4(B_2-B_3)(B_7−3B_8+3B_9 –B_{10})/B_1^2 + 6(B_2-B_3)^2 (B_4−2B_5+B_6)/B_1^3-3(B_2−B_3)^4/B_1^4]);
$M_{2,2}$:= evalf(b^{-2} [(C_4−2C_5+C_6)/C_1 − (C_2−C_3)^2 /C_1^2]);
$M_{3,2}$:= evalf(b^{-3} [(C_7−3C_8 +3C_9-C_{10})/C_1− 3(C_2-C_3)(C_4−2C_5+C_6)/C_1^2 +2 (C_2−C_3)^3/C_1^3]);
$M_{4,2}$:= evalf(b^4 [(C_{11}−4C_{12} +6C_{13}-4C_{14} +C_{15})/C_1− 4(C_2-C_3)(C_7−3C_8+3C_9 –C_{10})/C_1^2 + 6(C_2-C_3)^2 (C_4−2C_5+C_6)/C_1^3-3(C_2−C_3)^4/C_1^4]);
$M_{2,3}$:= evalf(b^{-2} [(D_4−2D_5+D_6)/D_1 − (D_2−D_3)^2 /D_1^2]);
$M_{3,3}$:= evalf(b^{-3} [(D_7−3D_8 +3D_9-D_{10})/D_1− 3(D_2-D_3)(D_4−2D_5+D_6)/D_1^2 +2 (D_2−D_3)^3/D_1^3]);

$M_{4,3}$:= evalf(b^4 [$(D_{11}-4D_{12}+6D_{13}-4D_{14}+D_{15})/D_1-4(D_2-D_3)(D_7-3D_8+3D_9-D_{10})/D_1^2$ +
$6(D_2-D_3)^2(D_4-2D_5+D_6)/D_1^3-3(D_2-D_3)^4/D_1^4$]);
α_1 :=H; α_2 :=$M_{2,1}+4M_{2,2}+M_{2,3}$; α_3 := evalf(($M_{3,3}-8M_{3,2}+M_{3,1}$)/ $\alpha_2^{1.5)}$;
α_1 := evalf($M_{4,1}+M_{4,3}+10M_{4,2}+24M_{2,3}M_{2,2}+24M_{2,2}M_{2,1}+6M_{2,3}M_{2,1}+6M_{2,2}^2/\alpha_2^2$);
Ah := [α_1, α_2, α_3, α_4]; end:

Programme-P2

\# Procedure to determine lambdas and the percentile points from sample Ah-values
\#Function: Findlambdas
\#Purpose: Estimation of GLD parameters by Newton's approx.
\#Arguments: Ah–list of a1, a2, a3, a4;
\# I3, I4–Initial approx. of λ3 and λ4
Findlambdas := Proc(Ah::list, I3::Numeric, I4::Numeric)
Local A, B, C, D1, D2, D, α1, α2, α3, α4, F, a1, a2, a3, a4, V, J, err3, err4, Fk, Jk, Y, Eq3, Eq4, A1, A2, L, FirstL, SecondL, l, R1, R2, R3, R4, R5, R6, R7, R8;
with(linalg, vector, matrix, jacobian, linsolve):
a1 := 0; a2 := 1; a3 := evalf(A_h[3]); a4 := evalf(A_h[4]); L3 := I3; L4 := I4;
A := [1/(1+ λ_3)]-[1/(1+ λ_4)];
B := [1/(1+ 2*λ_3)] + [1/(1+ 2*λ_4)]− 2* Beta(1 + λ3, 1 + λ4);
C := [1/(1+ 3*λ_3)]-[1/(1+ 3*λ_4)] + 3* Beta(1 + 2*λ3, 1 + λ4)+ 3* Beta(1 + λ3, 1 + 2* λ4);
D1 := [1/(1+ 4*λ_3)] + [1/(1+ 4*λ_4)] + 6* Beta(1 + 2*λ3, 1 +2* λ4);
D2 := −4 * Beta(1 + 3 * λ3, 1 + λ4) − 4 * Beta(1 + λ3, 1 + 3* λ4); D := D1 + D2;
α1 := λ1 + A/ λ2; α2 := abs(B − A^2)/ $\lambda 2^2$;
α3 := (C−3*A*B+2*$A^{3)}$/abs((B−$A^2)^{3/2}$) ;
α4 := (d−4*A*C+6*B*A^2−3*A^4)/(B−$A^2)^2$;
Eq3 := α3 − a3; Eq4 := α4 − a4;
F := vector([Eq3,Eq4]); V := vector([λ3, λ4]);
j := evalf(jacobian(F, V)); err3 := 1; err4 := 1;
while (err3 >.0001 or err4 >.0001) do
Fk := vector([evalf(subs(λ3 = L3, λ4 = L4,−Eq3)), evalf(subs(λ3 = L3, λ4 = L4,−Eq4))]);
Jk := matrix([[subs(λ3 = L3, λ4 = L4, j[1, 1]), subs(λ3 = L3, λ4 = L4, j[1, 2])],
[subs(λ3 = L3, λ4 = L4, j[2, 1]), subs(λ3 = L3, λ4 = L4, j[2, 2])]]);
Y := linsolve(Jk, Fk);L3 := L3 + Y [1];
L3 := L3 + Y [1]; L4 := L4 + Y [2];
err3 := evalf(abs(subs(λ3 = L3, λ4 = L4,Eq3)));
err4 := evalf(abs(subs(λ3 = L3, λ4 = L4,Eq4)));
od;
print(L3, L4, err3, err4);
A1 := evalf(subs(λ 3 = L3, λ4 = L4,A));
A2 := evalf(subs(λ3 = L3, λ4 = L4,B));
L2 := abs(sqrt(((A2 – A1^2))/a2));
L1 := a1 − A1/L2;

FirstL := [L1, L2, L3, L4];
if L3 < 0 then SecondL := [−FirstL[1], FirstL[2], FirstL[4], FirstL[3]] else SecondL := FirstL fi;
if evalf(Ah[3]) < 0 then L := [−SecondL[1], SecondL[2], SecondL[4], SecondL[3]] else
L := SecondL fi ;
l := [L[1] * sqrt(Ah[2]) + Ah[1], (L[2])/(sqrt(Ah[2])), L[3], L[4]];
t0.05 := l[1] + $((0.05)^{l[3]}–(0.95)^{l[4]})$/l[2] ; t0.95 := l[1] + $((0.95)^{l[3]}–(0.05)^{l[4]})$/l[2] ;
t0.025 := l[1] + $((0.025)^{l[3]}–(0.975)^{l[4]})$/l[2] ; t0.975 := l[1] + $((0.975)^{l[3]}–(0.025)^{l[4]})$/l[2] ;
t0.01 := l[1] + $((0.01)^{l[3]}–(0.99)^{l[4]})$/l[2] ; t0.99 := l[1] + $((0.99)^{l[3]}–(0.01)^{l[4]})$/l[2] ;
t0.005 := l[1] + $((0.005)^{l[3]}–(0.995)^{l[4]})$/l[2] ; t0.995 := l[1] +$((0.995)^{l[3]}–(0.005)^{l[4]})$/l[2] ;
end:

Programme-P3

Procedure to determine the power values of the test;
#Function: Power via iteration;
#Purpose: Compute power values of the test;
#Arguments: L–list of lambda values of the test;
#S_0, S_1– skewness values under H0, H1;
#P0–initial approx. of power;
#K–critical point of the test;
FindPower:= Proc(L:: list,Q0::Numeric, Q1::Numeric, P0::Numeric, K::Numeric)
Local l_1,Q,E,err,P,p;
l_1 := evalf(L[1] + $S_1 - S_0$); Q := l_1 + $(p^{L[3]} - (1 - p)^{L[4]})$/L[2];
P := P0; E := K − Q; err := 0.5; while (err >.00001) do P := P +.0002;
err:=evalf(subs(p = P,E)); od;
if the test is lower tailed then print(P,err) else print(1-P,err) fi; end:

References

[1] Balakrishnan, N. and Rao, C. R. (1998). Order Statistics Theory and Methods, Elsevier, NewYork.

[2] D'Agostino, R. B. (1971). An omnibus test test of normality for moderate and large size samples. Biometrika, 58(2), 341-348

[3] D'Agostino, R. B. and Pearson, E. S. (1973). Tests for departure from normality. Empirical results for the distributions of b2 and b1. Biometrika, 60, 613-622.

[4] David, H. A. (1981). Order Statistics, John Wiley & Sons, 2nd Edn., NewYork.

[5] David, H. A. and Nagaraja, H. N. (2003). Order Statistics, John Wiley & Sons, 3rd Edn., NewYork.

[6] Downton, F. (1966). Linear estimates with polynomial coefficients. Biometrika, 53, 129-141.

[7] Gupta, M. K. (1967). An asymptotically non-parametric test of symmetry. The Annals of Mathematical Statistics, 38, 849-866.

[8] Harter, H. L. and Balakrishnan, N. (1996). C.R.C Handbook of Tables for the Use of Order Staistics in Estimation, CRC Press, Boca Raton, FL.

[9] Karian, Z. A. and Dudewicz, E. J. (2000). Fitting Statistical Distributions-The Generalised Lambda Distribution and Generalised Bootstrap Methods, CRC press, NewYork.

[10] Kendall, M. G. and Stewart, A. (1977). The Advanced Theory of Statistics, Fourth ed., Vol.1 Griffin, London.

[11] King, R. and MacGillivary, H. L. (1999). A Starship Estimation for the Generalised Lambda Distributions, Australian and New-Zealand Journal of Statistics, 41(3), 353-374.

[12] Kumaran, M. and Beena, K. K. (2005). On Moments of Order Statistics from Generalised Lambda Distribution and Some Applications, Journal of Kerala Statistical Association, 16,20-30.

[13] Ozturk, A. and Dale, R. (1985); Least Squares Estimation of the Parameters of the Generalised Lambda Distribution, Technometrics, 27(1), 81-84.

[14] Pearson, E. S. (1965). Tables of the percentage points of b1 and b2 in normal samples : a round off. Biometrika, 52, 282-285.

[15] Pepper, J. (1932). The sampling distribution of the third moment coefficient-an experiment. Biometrika, 15(5), 55-64.

[16] Ramberg, J. and Schmeiser, B. (1974). An Approximate Method for Generating Asymmetric Random Variables, Communications of the ACM,17, 78-82.

[17] Ramberg, J. Tadikamalla, P. Dudewicz, E. J. and Mykytka, E. (1979). A Probability Distribution and Its Uses in Fitting Data, Technometrics, 21(2), 201-214.

[18] Reiss, R. D. (1989). Approximate Distributions of Order Statistics With Application to Nonparametric Statistics, Springer-Verlag, NewYork.

[19] Ross, S. M. (1987). Introduction to Probability and Statistics for Engineers and Scientists. John Wiley & Sons, New York.

[20] Shapiro, S. S. and Wilk, M. B. (1965). An analysis of variance test for normality (complete samples). Biometrika, 52, 591-611.