JRL of the Faculty of Commerce for Scientific Research.

Jan 2018,Vol. 55 No. 1

Prof.Dr.Osama Abdelaziz Hussien
Department Of Statistics
Faculty Of Commerce
Alexandria University

Robust Goodness-Of-Fit Tests Based On The TL-Moments

Abstract

In this paper we propose goodness-of-fit test based on the TL-moments. Elamir and Seheult (2003) introduced the trimmed L-moments (TL-moments) as a "robust "generalization of the method of L-moments. They are well defined even if L-moments do not exist. Their sample variance and covariance can be obtained in closed form, Elamir and Seheult (2004). A distribution for which the trimmed L-moments exist is characterized by those trimmed L-moments, Hosking (2007). Several goodness-of-fits tests were constructed based on "adaptive" proportions of trimming and different forms for the test statistic. Robustness of validity and robustness of efficiency will be discussed and compared to other robust goodness-of-fit e.g. Brys et al. (2008). Applications to real data will be conducted

Keywords: L-moments, TL-moments, LQ-moments, influence function, gross error sensitivity, breakdown point.

AMS Subject Classification : 62F35, 62F12, 62G35

1-Introduction

The frequency of occurrence of extreme events like high floods, drought, or heavy rainfall corresponds always with extraordinary observations that affect drastically the appropriate choice of the probability distribution and the efficiency of the classical estimators of its parameters, Hosking and Wallis (1993) . Even when a good fitted distribution can be found, there is no guarantee that future values will match those of the past, particularly when the data arise from a physical process that can give rise to occasional outlying values. One approach to decide the appropriate model is to fit a variety of plausible distributions and select the model with the smallest p-value of some goodness-of-fit test. The classical goodness of fit tests depend on estimating the parameters of the null distribution by the MLE which is not robust (its efficiency reduced drastically by small deviations from the assumed model). This leads to lower power of the test and different p-values especially for skewed distributions; Ronchetti (1997). Hertiteir and Ronchetti (1994) showed that finding the test which maximizes the asymptotic power subject to a bound on the level and power influence function is equivalent to finding an estimator T which minimizes the asymptotic variance subject to a bound on its self-standardized influence function. Michael and Schucany (1985) proved that the Komogrov-Smirnov statistic and the Cramer-von Mises statistics have a bounded influence function. That is, the significance level of the test is robust to small amount of contamination in the model.

In regional frequency analysis the L-moments and the TL-moments estimators are used extensively to estimate the population parameters. The L-moments are defined as linear combinations of expected values of order statistics of an absoulutly continous random variable, (Hosking 1990).The advantages of L-moments over classical moments are: able to charactrize a wider range of distributions; and the L-moments estimators are “more robust to the presence of outliers”. Elamir and Seheult (2003) introduced an extension of L-moments called TL-moments. TL-moments estimators are assumed to be “more robust against outliers” than L-moments estimators.The population TL-moment may be well defined where the corresponding population L-moment does not exist. Hussien (2011) derives the influence functions and the breakdown points of the L-moments and the TL-moments estimators. Unlike the TL-mo-ments estimators the L-moments esti-mators are not locally nor globaly rob-ust.The TL-moments estimators are not only robust but it can be adapted ac-cording to the shape of the data by varying the proportions of trimming (symmetric trimming for heavy tailed shapes and asymmetric trimming for skewed shapes).

Dalen (1987) showed that the classical moment ratios are bounded and cannot attain the full range of values available to the population skewness and kurtosis. The sample TL-skewness TL-kurtosis ratios graph are visual tool to goodness of fit analog to the, non robust, goodness of fit tools based on sample skewness and kurtosis. Moreover, it can take any of the feasible values of the population TL-moment ratios, Hosking (2007). Thus the TL- skewness TL-kurtosis ratios graph would play the role of an exploratory stage and a goodness of fit test would play the role of a confirmatory stage by choosing the distribution with the smallest p-value among the candidate distributions in the first stage.

Using a robust estimator (like the TL-moments estimators) does not guarantee the robustness of the goodness of fit test depends on these estimators. For a test procedure to be robust the level of a test should be stable under small, arbitrary departures from the null hypothesis (robustness of validity). Also, the test should still have a good power under small arbitrary departures from specified alternatives (robustness of efficiency), Huber and Ronchetti (2009). A robust (optimal bounded- influence) tests can be obtained by finding a test in a given class that maximizes the asymptotic power at the model subject to a bound on the level and power influence functions. For parametric tests the Wald-type tests and the score-type tests based on M-estimators are proved to be robust, Hertiteir and Ronchetti (1994). This paper presents Wald-type tests and score-type tests for goodness of fit based on TL-estimators and discuss its robustness properties.

The rest of the paper is organized as follows; in Section 2 discuss the robustness properties of goodness of fit tests and derives finite sample formula for influence function of the test statistic. Section 3 discusses the robustness properties of the TL-estimators. In Section 4 we derive the optimal bounded influence goodness of fit tests based on the TL-moments and study its properties. Some Monte Carlo simulation results and a study of the River Nile data are provided in Section 5.

2. Robust Goodness of Fit Tests

Consider $X_1, X_2, \ldots, X_n$ iid with cumulative distribution function F_θ. Let T(.) be a weakly continuous functional with values in P^p, where $T_n(x_1,x_2,\ldots,x_n) = T(F_n)$ is an estimator of $\theta \in \Omega$, an open convex subset of P^p, and F_n is the empirical cumulative distribution function. The functional T is said to be Fisher consistent for θ if $T(F_\theta)=\theta$ for all $\theta \in \Omega$.(see Rao, 1975 page 345). The influence function of a statistical functional T at a distribution F is defined by

$$IF(x;T,F) = \lim_{\epsilon \to 0^+} \frac{T[(1-\epsilon)F+\epsilon\delta_x]-T(F)}{\epsilon} \qquad (1)$$

rovided that the limit exist, where δ_x is a degenerate density function that puts all its mass at the point x. Under regularity conditions sufficient for the von Mises expansion

$$T(F_n) - T(F) = \int_I IF(x;T,F)dF_n(x) + o_p(n^{-0.5}) \quad (2)$$

to exist, we have $\int IF(x;T,F)dF(x) = 0$ and

$$\sqrt{n}\left(T(F_n) - T(FP)\right) \xrightarrow{d} N(0, AV(T,F)), \quad (3)$$

with

$$AV(T,F) = \int IF(x;T,F)\, IF(x;T,F)^t dF(x), \quad (4)$$

See, Hampel et al. 1986.

An estimator T(F) is locally robust if its influence function is bounded. For a positive definite matrix B, the gross error sensitivity of a statistical functional T is given by

$$\gamma(T,F,B) = \sup_{x \in I}[IF(x;T,F)^t B\, IF(x;T,F)]^{\frac{1}{2}} \quad (5)$$

The gross error sensitivity provides an upper bound for the asymptotic bias of $T(F_n)$ over a sufficiently small ε-contamination neighborhood of the assumed model. When $B=I_p$ one gets the so-called unstandardized gross error sensitivity (Hampel et al., 1986). When the norm generated by the inverse of the asymptotic variance matrix, $B = AV(T,F)^{-1}$ it produces the self-standardized gross error sensitivity (Krasker and Welch, 1982). The breakdown point of an estimator describes up to what distance from the model distribution the estimator still gives some relevant information. An estimator with a breakdown point close to 0.5 is to be global robust.

The computation of the influence functions needs tedious numerical integrations for each probability distribution. Instead one may compute the discredized form of the influence function given by

$$SC(T_n, y) = n[T_n(x_1, x_2, \ldots, x_{n-1}, y) - T_{n-1}(x_1, x_2, \ldots, x_{n-1})] \qquad (6)$$

When $T_n(x_1, x_2, \ldots, x_n) = T(F_n)$ for any n and any sample $(x_1, x_2, \ldots, x_n)$, then

$$SC(T_n, y) = \left[T\left(\left(1-\tfrac{1}{n}\right)F_{n-1} + \tfrac{1}{n}\delta_y\right) - T(F_{n-1})\right] / \tfrac{1}{n}$$

, and

$$\lim_{n\to\infty} SC(T_n, y) = IF(Y, T, F)$$

,see Hampel et al. (1986). The number

$$S(T_n) = sup_y|SC(Tn, y)|$$

is called a sensitivity of the functional T_n to an additional observation, Jureckova and Picek (2006).

The two fundamental goals in robust testing are:(1) the level of the test should be stable under small, arbitrary departures from the null hypothesis (robustness of validity), and (2) the test should retain a good power under small, arbitrary departures from specified alternatives (robustness of efficiency). Many classical inferential procedures do not satisfy these criteria. The non-robustness of the test is due simultaneously to

1- The non-robustness of the parameter estimator ,

2- The non-robustness of the test statistic.

Consider a parametric model, $\{F_\theta\}$ where θ is a real parameter Let $H_0 : \theta = \theta_0$ be the null hypothesis and $\theta_n = \theta_0 + \frac{\Delta}{\sqrt{n}}$ a sequence of alternatives. Assume the test statistic T_n is Fisher consistent and the regularity conditions that assures (2), (3) and (4) above are satisfied. Consider a neighborhood of distributions $F_{\varepsilon,n}=\left(1-\frac{\varepsilon}{\sqrt{n}}\right)F_{\theta_0}+\frac{\varepsilon}{\sqrt{n}}G$ where G is an arbitrary distribution. One can view the asymptotic level α of the test as a functional of a distribution in the neighborhood. Then by a von Mises expansion of α around F_{θ_0}, where $\alpha(F_{\theta_0})=\alpha_0$,the nominal level of the test, the asymptotic level and (similarly) the asymptotic power under contamination can be expressed as

$$\lim_{n\to\infty}\alpha\left(F_{\varepsilon,\theta_0,n}\right)=\alpha_0+\varepsilon\int LIF\left(x;T,F_{\theta_0}\right)dG(x)+o(\varepsilon) \quad (7)$$

$$\lim_{n\to\infty}\beta\left(F_{\varepsilon,\theta_n,n}\right)=\beta_0+\varepsilon\int PIF\left(x;T,F_{\theta_0}\right)dG(x)+o(\varepsilon) \quad (8)$$

where

$$LIF\left(x;T,F_{\theta_0}\right)=\phi(\Phi^{-1}(1-\alpha_0))IF\left(x;T,F_{\theta_0}\right)/[AV(T,F_{\theta_0})]^{1/2} \quad , (9)$$

$$PIF\left(x;T,F_{\theta_0}\right)=\phi(\Phi^{-1}(1-\alpha_0)-\Delta\sqrt{E})IF\left(x;T,F_{\theta_0}\right)/[AV(T,F_{\theta_0})]^{1/2}, \quad (10)$$

$\alpha(F_{\theta_0})=\alpha_0$ is the nominal asymptotic level, $\beta_0=1-\Phi(\Phi^{-1}(1-\alpha_0)-\Delta\sqrt{E})$ is the nominal asymptotic power, $E=[\xi'(\theta_0)]^2/AV(T,F_{\theta_0})$ is Pitman's efficacy of the test, $\xi(\theta)=T(F_\theta)$, $AV(T,F_{\theta_0})=\int IF(x;T,F_{\theta_0})^2 dF_{\theta_0}(x)$ is the asymptotic variance of T, and $\Phi^{-1}(1-\alpha_0)$ is the 1-α_0 quantile of the standard normal distribution Φ and ϕ is its density. It follows that the level influence function (9) and power influence function (10) are proportional to the self-standardized influence function of the test statistic $IF_s\left(x;T,F_{\theta_0}\right)=IF\left(x;T,F_{\theta_0}\right)/[AV(T,F_{\theta_0})]^{1/2}$,Huber and Ronchetti (2009). Thus, one can obtain the maximal level and the minimal power over the neighborhood as

$$\sup_G \text{as level} \cong \alpha_0+\varepsilon\phi(\Phi^{-1}(1-\alpha_0))\sup_x\left(IF\left(x;T,F_{\theta_0}\right)/[AV(T,F_{\theta_0})]^{\frac{1}{2}}\right)$$

$$\inf_G \text{as power} \cong \beta_0+\varepsilon\phi(\Phi^{-1}(1-\alpha_0)-\Delta\sqrt{E})\inf_x\left(IF\left(x;T,F_{\theta_0}\right)/[AV(T,F_{\theta_0})]^{\frac{1}{2}}\right)$$

Therefore,bounding the influence function of the test statistic T from above will ensure robustness of validity and bounding it from below will ensure robustness of efficiency. This will not generally guarantee that the level and the power will remain stable in the presence of large deviations. The effect of large deviations is described by the breakdown point. A finite sample definition of the breakdown point of a test was introduced by Ylvisaker (1977). Consider a test with critical region ($T_n \geq c_n$).The resistance to acceptance ϵ_a^* [resistance to rejection ϵ_r^*] of the test is defined as the smallest proportion m/n for which there are values $x_1, x_2, \ldots, x_m$ in the sample with $T_n < c_n$ [$T_n \geq c_n$].

In the multivariate case and for general parametric models, the classical theory provides three asymptotically equivalent tests, Wald, score, and likelihood ratio test, which are symptotically uniformly most powerful with respect to a sequence of contiguous alternatives. If the parameters of the model are estimated by a robust estimator such as an M-estimator T_n defined by the estimating equation $\sum_{i=1}^{n} \psi(x_i, T_n) = 0$, then a robust form of the three classical tests can be constructed by replacing the score function of the model by the function ψ. This leads to formulas similar to (7) and (8) and to optimal bounded influence tests; Heritier and Ron-chetti (1994).

Bounded influence robust goodness of fit tests where proposed by Vectoria-Feser (1993). For the non-nested hypotheses: H_o: F=F_{α}^{0} vs. H_1: F= F_{β}^{1}, with α and β being two sets of parameters, Vectoria-Feser (1997) showed that the LIF for the Cox test statistics, Cox (1963), is unbounded. Vectoria-Feser (1997) therefore proposed a robust Cox-type test statistic based on the optimal B-robust M-estimator presented by Hampel et al. (1986). However this test procedure does not have a high breakdown point especially when the dimension of the parameters is large.This means the test procedure will breakdown when a cluster of outliers is presentin thedata;Vectoria-Feser (2000).

2.1.P-value Influence function

The level and power influence functions depend on the chosen error probabilities. Alternatively, one might consider the influence of outliers on the test statistic or some function of it, as the p-value. The P-value or observed significance level is used to assess the strength of the evidence against a null hypothesis H_0. Assume the test statistic T_n has been chosen so that large values indicate significant departures from H_0. Lambert and Hall (1982) showed that the P- value is asymptotically log- normal under an alternative F_θ if the test statistic T_n, is asymptotically normal and if the tail of a cdf of its sampling distribution under the alternative satisfying weak regularity conditions. This shows that the mean, which is half the Bahadur slope $c(F_\theta)$, and the standard deviation $\tau(\theta)$ of the asymptotic distribution of the log transformed P-value permit approximation of both the level and power of the test. The influence function for $-n^{-1} \log P_n$ defined by

$$\mathbf{IF(x;P,F)} = \lim_{\epsilon \to 0^+} \frac{c[(1-\epsilon)F_\theta + \epsilon\delta_x] - c(F_\theta)}{\epsilon} \quad (11)$$

describes the extent of the error in the observed significance level relative to its size, Lambert (1981). If $c(F_\theta)$ has a finite first derivative c' in a neighborhood of $T(F_\theta)$ and IF(x;T,F) then

$$\mathbf{IF\,(x;P,F_\theta) = c'(F_\theta) IF(x;T,F_\theta)} \quad (12)$$

Definition: P-value sensitivity curve

The p-value influence function evaluates the effect of outliers on the test under the alternative. A corresponding p-value sensitivity curve is defined by

SC(P$_n$, y) =

$$\left[-\frac{1}{n}\log P_n\left(\left(1-\frac{k_n}{n}\right)F_{n-1}+\frac{k_n}{n}\delta_y\right)+\frac{1}{n-k_n}\log P_n\left(F_{n-1}\right)\right]/k_n/n$$

Lambert (1981) showed that

$$\lim_{n\to\infty} SC(P_n, y) = IF(Y;, P, F_\theta)$$

As an alternative to the level influence function using the p-value approach we define

$$SC(P_n, y|H_0)=n[\frac{1}{n} P_n (x_1, x_2, \ldots, x_{n-1}, y|H_0) - \frac{1}{n-1} P_n (x_1, x_2, \ldots, x_{n-1}|H_0)] \quad (13)$$

It measures the effect of adding outliers on the p-value.The slope of this curve is an indication of the validity robustness of the test. If SC(P$_n$, y|H_0) < α this is an indication that the test breakdown. Thus we can estimate the breakdown point of the test by

(number of y's such that SC(P$_n$, y|H_0) > α).

2.2 Moments based goodness of fit tests

Under regularity conditions, the infinite set of moments (when these exist) characterize a distribution; thus, it is appealing to use the first s sample moments, or functions of these, as test statistics for goodness of fit. For example, the sample skewness $b_1 = m_3^2/m_2^2$ and the sample kurtosis $b_2 = m_4/m_2^2$, where, $m_r = \frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^r$ have long been used to test for normality.Gurland and Dahiya (1970) and Dahiya and Gurland (1972) developed a procedure for goodness of fit based on the sample moments. The method depends on being able to find both a vector function E of length s of the first s population moments and a parameterization θ of the distribution, where θ has length q, q<s, such that E is linear in θ. The test statistic has an asymptotic χ_t^2 distribution, t=s-q. The influences functions for the central moments are unbounded, Sehđrlđoğlu (2010). Thus they are not locally or globally robust. Moreover its breakdown points are zero. Thus, the goodness of fit tests based on the sample moments are not robust also.

The above tests can be viewed as a special case of the following generalization. Let $w=(w_1, w_2, \ldots, w_s)'$ be an (unbiased) estimators of $\Lambda = (\Lambda_1, \Lambda_2, \ldots \Lambda_s)'$, such that

$$\sqrt{n}\, w \xrightarrow{D} N(\Lambda, \Sigma)$$

then , under H_0 the generalized test statistic

$$T=n(w-\Lambda)'\Sigma^{-1}(w-\Lambda)\approx\chi_s^2 \quad (14)$$

Brys, et al. (2008) propose several goodness-of-fit tests based on robust measures of skewness and tail weight. The robust skewness measure, proposed by Brys et al. (2003) is defined as

$MC(F) = \text{med}_{x_1<m_F<x_2}\, h(x_1, x_2)$

with x_1 and x_2 sampled from F, $m_F = F^{-1}(0.5)$ and the kernel function h given by

$$h(x_1,x_2) = \frac{(x_j-m_F)-(m_F-x_i)}{x_j-x_i}$$

This estimator has a breakdown value of 0.25 and a bounded influence function, Brys et al. (2004).

As robust measures of the left and right weight, Brys et al. (2005) applied the MC to the left and right half of the sample respectively:

LMC(F) =-MC(x < m_F) and RMC(F) = MC(x > m_F),

yielding a breakdown value of 0.125. These measures can be computed even when the finite moments does not exist.

Brys et al. (2008) studied the pr-operties of the MC-LR test of the form (11) where s=3, w_1=MC, w_2=LMC, and w_3=RMC. Also, they study the test of Moors et al. (1996) where

$$w_1 = \frac{F^{-1}(0.75)+2F^{-1}(0.5)-F^{-1}(0.25)}{F^{-1}(0.75)-F^{-1}(0.25)}$$

and

$$w_2 = \frac{F^{-1}(0.875)-F^{-1}(0.625)+F^{-1}(0.375)-F^{-1}(0.125)}{F^{-1}(0.75)-F^{-1}(0.25)}$$

are robust measures of skewness and kurtosis. Both tests has a breakdown value of 0.125. They did not derive the level influence curves of the above te-sts,but a simulation study showed that both MOORS and MC-LR are not high-ly influenced by outliers.

3.The TL-moments estimators

Let X be a real values continuous random variable with cumulative distri-bution function cdf F(x), and let $X_{1:n} \leq X_{2:n} \leq \ldots \leq X_{n:n}$ be the order statistics of a random sample of size n drawn from the distribution of X.The L-moments of X are expectations of certain co-mbi-nations of order statistics defined by

$$\lambda_k = k^{-1}\sum_{j=0}^{k-1}(-1)^j\binom{k-1}{j}E\left(X_{k-j:k}\right), \quad k=1,2,\ldots \tag{15}$$

Substituting into (1) the standard expression of expected value of order statistics

$$E\left(X_{r:n}^k\right)=\frac{n!}{(r-1)!(n-r)!}\int_{-\infty}^{\infty} y^k[F_X(y)]^{r-1}[1-F_X(y)]^{n-r}f_X(y)dy$$

yields the classical L-functional pre-sentation of the distribution L-moments given by Hosking (1990) as

$$\lambda_k(F)=\int_0^1 F^{-1}(u)P_{k-1}^*(u)du \tag{16}$$

where

$$F^{-1}(u) = inf\{x: F(x) \geq u\}$$

$$,P_r^*(u) = \sum_{j=0}^{r} p_{r,j}^* u^j \tag{17}$$

and

$$p_{r,j}^* = (-1)^{r-j}\binom{r}{j}\binom{j+r}{j}$$

$P_r^*(u)$ is the rth shifted Legender poly-nomial related to the usual Legender polynomial by the relation $P_r^*(u) = P_r(2u-1)$.

Hosking (1990) showed that the estima-tor of λ_k

$$\lambda_k(F_n)=\int_0^1 F_n^{-1}(u)P_{k-1}^*(u)du \tag{18}$$

$=\sum_{i=1}^{n} c_{ni}^k X_{i:n}$,where $c_{ni}^k = \int_{\frac{i-1}{n}}^{\frac{i}{n}} P_{k-1}^*(u)du$ is unbiased.

Under suitable regularity condi-tions, Hosking (1990) proved that $\sqrt{n}(\hat{\lambda}_k - \lambda_k)$, k=1,2,…,m converges in distribution to the multivariate normal distribution.

Elamir and Seheult (2003) intro-duced the trimmed L-moments (TL-moments) as a “robust “generalization of the method of L-moments. The TL-moments assigns zero weight to ex-treme observations, and they are well defined even if L-moments do not exist. The k^{th} population TL-moment with trimming proportion (s,t) is defined by:

$$\lambda_k^{s,t}=\frac{1}{k}\sum_{j=0}^{k-1}(-1)^j\binom{k-1}{j}E(X_{k+s-j:k+t+s}) \quad (19)$$

The TL-moments can be written as

$$\lambda_k^{s,t}=\int_0^1 F^{-1}(u)\, P_k^{s,t}(u)\,du \quad (20)$$

where

$$P_k^{s,t}(u)=\frac{(k-1)!\,\Gamma(k+s+t+1)}{k!\,\Gamma(k+s)\,\Gamma(k+t)}\sum_{j=0}^{k-1}(-1)^{k-1-j}\binom{k-1+s}{k-1-j}\binom{k-1+t}{j}u^{j+s}(1-u)^{k-1-j+s}$$

An unbiased kth sample TL-moment with trimming proportion is given by Elamir and Seheult (2003) as

$$l_k^{s,t}=\frac{1}{k\binom{n}{k+s+t}}\sum_{i=s+1}^{n-t} w_k^{s,t}(i)\,x_{i:n} \quad (21)$$

with

$$w_k^{s,t}(i)=\frac{1}{\binom{n}{k+s+t}}\sum_{j=0}^{k-1}(-1)^j\binom{k-1}{j}\binom{i-1}{k+s-j-1}\binom{n-i}{t+j} \quad (22)$$

Under suitable regularity conditions Hosking (2007) proved that $\sqrt{n}\left(\ell_k^{s,t}-\lambda_k^{s,t}\right), k=1,2,\ldots,m$ converges in distribution to the multivariate normal distribution $N(0,\Lambda^{s,t})$, where the elements $\Lambda_{uv}^{s,t}$ (u,v=1,2,…,m) of $\Lambda^{s,t}$ is given by

$$\Lambda_{uv}^{s,t}=\int\int_{x<y}\{P_u^{s,t}(F(x))PH_v^{s,t}(F(y)) + P_v^{s,t}(F(x))P_u^{s,t}(F(y))\,F(x)(1-F(y)\}dxdy \quad (23)$$

The exact variances and covariance of the sample TL-moments are given by Elamir and Seheult (2003).The R backage lmomco, Asquith (2011), computes the sample TLmoments and the theoretical TLmoments for several distributions. Hosking (2007) (Theorem2) proved that a distribution for which the TL-moment exist is characterized by those TL-moments.

Hussien (2011) showed that the kth TL-moment estimator is locally and globally robust while the kth L-moment estimator is not.

Lemma 1; Hussien (2011)

1- The influence function of the kth L-moment estimator is unbounded, its gross error sensitivity is infinity, and its finite breakdown point is zero.

2- The influence function of the kth TL-moment estimator is given by

$$IF_k\left(z,\lambda_k^{\alpha,\beta},F\right)=\begin{cases}\int_{F^{-1}(\alpha)}^{F^{-1}(1-\beta)}(F(v)-1)w^{\alpha,\beta}(v)dv & z\le F^{-1}(\alpha)\\ \int_{F^{-1}(\alpha)}^{z}F(v)\,w^{\alpha,\beta}(v)dv+\int_{z}^{F^{-1}(1-\beta)}(F(v)-1)w^{\alpha,\beta}(v)dv & F^{-1}(\alpha)\le z\le F^{-1}(1-\beta)\\ \int_{F^{-1}(\alpha)}^{F^{-1}(1-\beta)}F(v)\,w^{\alpha,\beta}(v)dv & z>F^{-1}(1-\beta)\end{cases} \quad (24)$$

3- The gross error sensitivity of $\lambda_k^{\alpha,\beta}$ is

$$\gamma^*=\left|\int_{F^{-1}(\alpha)}^{F^{-1}(1-\beta)}F(v)\,w^{\alpha,\beta}(v)dv\right|<\infty \quad (25)$$

Hosking (1990) showed that the L-moment ratios, defined by

$\tau_k=\lambda_k/\lambda_2$ k=3,4

are appropriate measures of the population skewness and kortosis. Hosking and Wallis (1997) showed that the L-moment ratio estimators are nearly unbiased for all sample sizes and all distributions. Values of τ_3 and τ_4 are plotted to yield an *L*-moment ratio diagram. Parametric families of distributions may occupy points, lines or regions on the graph and the sample *L*-moments ratios of data sets can be plotted as points. Thus, in a model choice process the L-moment diagram would play the role of an exploratory stage, where candidate distributions are selected, and a goodness of fit test would play the role of a confirmatory stage. Graph 1 gives the SC($\hat{\tau}_3, y$) and SC ($\hat{\tau}_4, y$) for a random sample of size 100 from a standard normal distribution.This shows that influence functions of τ_3 and τ_4 are unbounded. This is also true for any absolutely continuous distribution and for all sample sizes. The breakdown points for the L-moment ratio estimators are zero, Hussien (2011).

The TL-moment ratios are defined by

$$\tau_k^{s,t} = \lambda_k^{s,t} / \lambda_2^{s,t} \quad k=3,4$$

Hosking (2007) showed that sample trimmed *L*-moment ratios provide an effective way of distinguishing between heavy tailed distributions. The trimmed *L*-moment ratio diagram covers a wider range of distributions; for example the Cauchy distribution. Graph 2 gives the SC($\hat{\tau}_3^{1,5}, y$) and SC($\hat{\tau}_4^{1,5}, y$)for a random sample of size 100 from a standard normal distribution. The graph shows that the influence functions are bounded.The same results hold for many probability distributions we considered e.g. gamma with different shape parameter, the lognormal, and the extreme value distribution. One would prefer to use the TL-moment ratio diagram in the exploratory stage, since its influence functions are bounded. Elamir (2010) derive an optimal choice for the amount of trimming from known distributions based on the minimum sum of the absolute value of the errors between the sample quantile function and its TL-moments representation

$$e(F_n)=F_n^{-1}(u)-\sum_{k=0}^{\infty}\frac{(k+1)(2k+s+t+1)}{k+s+t+1}l_{k+1}^{s,t}T_k^{s,t}(F_n) \quad (26)$$

where

$$T_k^{s,t}(F_n)=\sum_{j=0}^{k}(-1)^{k-j}\binom{k+t}{k-j}\binom{k+s}{j}F_n^j(1-F_n)^{k-j}$$

Accordingly, one can determine the appropriate amounts of trimming (s^*,t^*) as the values that minimizes $\sum|e(F_n)|$, then use the TL-moment ratio diagram with (s^*,t^*) to determine the candidate distributions. The confirmatory stage should use a robust goodness of fit test. Thus we suggest constructing a robust goodness of fit test that based on the TL-moments. The next section proposes two test statistics based on the TL-moments and prove its robustness properties.

4. Description of the tests

Let $X_1,X_2,\ldots,X_n$ be a sequence of iid random variables with an underlying continuous cdf F. Consider the following hypotheses: H_{0s} : G(x) = F(x;θ_0) versus H_{1s} : G(x)≠ F(x;θ_0) where $\theta_0 \in \Omega \subseteq R^q$ is a specified value of the vector parameter θ.

Define $\lambda_0^{s,t} = (\lambda_1^{s,t},\lambda_2^{s,t},\lambda_3^{s,t},\ldots,\lambda_m^{s,t})^t$ to be the vector of the first four TL-moments computed for F(x, ;θ_0), i.e. under H_0. Also, let $\ell^{s,t} = (\ell_1^{s,t},\ell_2^{s,t},\ell_3^{s,t},\ldots,\ell_m^{s,t})^t$ be the sample four TL-moments.

Consider the generalized test statistics

$$T_1=(\ell^{s,t}-\lambda_0^{s,t})^t(\Lambda^{s,t})^{-1}(\ell^{s,t}-\lambda_0^{s,t}) \quad (27)$$

Lemma (2):

Under H_0 the test statistic

$$T_1= n(\ell^{s,t}-\lambda_0^{s,t})^t(\Lambda^{s,t})^{-1}(\ell^{s,t}-\lambda_0^{s,t})$$

has an approximate chi-square distribution with m df

Proof:

Hosking (2007) proved that $\sqrt{n}(\ell_k^{s,t} - \lambda_k^{s,t}), k = 1, 2, ..., m$ converges in distribution to the multivariate normal distribution $N(0, \Lambda^{s,t})$, where the elements $\Lambda_{uv}^{s,t}$ (u,v=1,2,…,m) of $\Lambda^{s,t}$ is given by (20). The result is immediate by (11).

Lemma (3)

Under the regularity conditions, necessary for the asymptotic normality of $\ell^{s,t}$, the p-value influence function of T_1 is given by $IF(x;P,F_\theta) = c'(F_\theta)\, IF(x;T_1, F_\theta)$

where

$c'(F_\theta)$ = and

$IF(x;T_1, F_\theta)$ =

Proof:

The p-value sensitivity curve of T_1 is given in Graph ( ).

This shows that …….

For the comparison we will graph the p-value sensitivity curves of the following test statistics:

$$T_{11} = n(\ell^{s,t} - \lambda_0^{s,t})'(\Lambda^{s,t})^{-1}(\ell^{s,t} - \lambda_0^{s,t})$$

s=t=1

$$T12 = n(\ell^{s,t} - \lambda_0^{s,t})'(\Lambda^{s,t})^{-1}(\ell^{s,t} - \lambda_0^{s,t})$$ **s=0, t=1**

$$T13 = n(\ell^{s,t} - \lambda_0^{s,t})'(\Lambda^{s,t})^{-1}(\ell^{s,t} - \lambda_0^{s,t})$$

s=0, t=3

$$T14 = n(\ell^{s,t} - \lambda_0^{s,t})'(\Lambda^{s,t})^{-1}(\ell^{s,t} - \lambda_0^{s,t})$$

s=0, t=5

$$T_2 = n(\ell - \lambda_0^{t})'\Lambda^{-1}(\ell - \lambda_0)$$

T_3 is the MOORS test

T_4 is the MC-LR test

$$T_5 = \sum |e(F_n)|$$

Graphs shows that

A suggested algorithm for robust goodness of fit

Step0: determine the appropriate proportion of trimming (if any) by the Box plot and solving $e(F_n)$

Step 1: draw the tl ratio diagram (with s and t determined in step 0)

Step 2: plot the sample tl ratios on the tl ratio diagram to select the candidate distributions

Step 3: apply the goodness of fit with test statistic T1 or T2 for all candidate distributions

Step 4: select the distribution with the highest p-value.

5. Example Application: Flood frequency distribution for the River Nile:

The River Nile data has the advantage of being long recorded through history (Yearly and monthly data available from 1870-1993). The yearly data is the maximum flood inflow at Aswan.

The Generalized Lambda Distribution (GLD) best fit procedure we use utilize TL-moments with different trimming proportions from each tail, and judge the best fit according to a goodness of fit test p-values.

Graph (3) below is a statistical summary of the data available (1870-1991) which constitutes a time-series of 122 observations. Presence of 2 outliers is clear from the graph (year 1913 drought and year 1878 high flood). The data is slightly right-skewed. It is apparent that classical parametric estimation methods would not be appropriate and more robust estimation methods are required.

The GLD best fit procedure:

1. Calculate first four sample estimates relevant to each estimation method:

The first four moments, first four L-moments , first four TL-moments(1), TL-moments(2), TL-moments(5), TL-moments(0,1), TL-moments(0,2), TL-moments(0,5), TL-moments(1,2), TL-moments(1,5) and the percentiles method.

2. Select suitable initial values for the optimization process, values of [0.1, 0.1, 0.1] were chosen as an initial estimate for the GLD parameters (η_2, η_3, η_4) for all methods.

An optimization process is started to find the best values of η_2, η_3, η_4

3. The output of the optimization process is the best attained values of the parameters (η_2, η_3, η_4).
4. Calculate the value of λ_1 given the optimized parameters. Generate 122 observations from the uniform distribution, using the estimates of (η_1, η_2, η_3, η_4), transform them into a GLD (η_1, η_2, η_3, η_4).

The Matlab procedures of Hussien and Afifi (2009) were used to conduct all the steps above.

5. Run a two-sample Kolmogorov-Smirnov test between the flood data and the GLD ($\hat{\eta}_1$, $\hat{\eta}_2$, $\hat{\eta}_3$, $\hat{\eta}_4$) generated data.
6. The method with the highest P-value for the Kolmogorov-Smirnov test is considered to give the best fit for the data.
7. Find the distribution approximated by the best GLD in step 6. Karian and Dudewicz (2000) give tables for these approximations for the method of moments only. The R-cran package “lmomco” draw the L-skewness L-kurtosis ratios graph. Many distributions are presented on the graph by point, line or area. You choose graphically the distribution nearest to the sample L-skewness L-kurtosis ratios. For this study we utilize this graph to decide the appropriate parent distribution. A simulation study is needed to draw new graphs for the TL-moments ratios for different proportions of trimming.

A summary of the results is attached in the following table.

Table (1) Estimates of the GLD parameters for the River Nile data.

	$\hat{\varepsilon}_1$	$\hat{\varepsilon}_2$	$\hat{\varepsilon}_3$	$\hat{\varepsilon}_4$	$\hat{\lambda}_1$	$\hat{\lambda}_2$	$\hat{\lambda}_3$	$\hat{\lambda}_4$	τ_3	τ_4	P-value
moments	90.6662	18.7911	0.5414	3.1565	79.3073	0.0097	0.0583	0.1977			0.6941
L-moments	90.6662	10.5085	1.3012	1.3310	78.8783	0.0079	0.0439	0.1570	5.5253	19.7622	0.8059
percentile	88.000	48.900	0.6355	0.4652	83.5744	-0.0183	-0.1104	-0.2147			0.4804
TL(1)	89.3650	5.5065	0.5912	0.4165	79.2051	0.0034	0.0178	0.0614	5.2314	18.0431	0.5016
TL(2)	88.8329	3.6952	0.3158	0.2098	79.9975	-0.0001	-0.00005	-0.0015	5.4489	16.7159	0.8351
TL(5)	88.3005	1.8158	0.0897	0.0446	79.9675	0.0000	0.0000	0.0001	5.2393	16.5533	0.9678
TL(0,1)	80.1576	6.9055	-0.0199	0.7946	80.2370	0.0008	0.0044	0.0130	5.6054	16.6799	0.9616
TL(0,2)	75.5539	5.5364	-0.4683	0.6203	80.2880	0.0000	0.0000	0.0001	5.5796	16.4911	0.5942
TL(0,5)	69.3671	4.1693	-0.8260	0.5545	80.9874	0.0000	0.0000	0.0000	5.7868	15.3429	0.4338
TL(1,2)	83.8585	4.1454	0.1657	0.2523	79.9093	0.0001	0.0006	0.002	5.3986	16.8062	0.9962
TL(1,5)	76.5144	2.7424	-0.1381	0.0824	89.2677	0.0308	0.4816	0.4144	15.6105	13.4325	0.4411

Results and comments:

1- The method of TL-moments(1,2) gave the best fit to be GLD (79. 9093, 0.0001, 0.0006, 0.002) with P-value =0.9962.

2- The method of TL-moments(0, 5) gave the worst fit to be GLD (80.9874, 0.0000, 0.0000, 0.000079) with P-value =0.4338.

3- The L-moment ratio diagram for Nile flood data is displayed in figure (5), with values of the sample τ_3 and τ_4 marked as a big circle in the middle of the figure. The figure suggests that this pair of values could coincide with either "the Rayleigh distribution" or "the Pearson type III distribution".

4- The L-moments ratio diagram is a graphical tool. To compare its result to the best fit procedure we need to do the following:

a- Construct a TL-moment(t_1,t_2) ratio diagram, so we compare the best TL- moments estimates obtained with the "nearest point" in the ratio diagram. The diagram should cover a wide class of different shapes.

b- Convert the best TL-estimates to the corresponding parameters of the distribution chosen by the diagram.

c- Generate a sample of the same size as the original data from the chosen distribution, F_{dgr} say, with the parameters estimated above.

d- Generate a sample of the same size as the original data from the GLD suggested by the best fit procedure, F_{GLD} say.

e- Construct a two-sample Kolmogorov-Smirnov test to test H_0: F_{dgr}= F_{GLD}. Construct another two-sample Kolmogorov-Smirnov test to test H_0: F_{dgr}=F_θ (the original distribution).

Unfortunately, new algorithms is needed to perform steps a and b.

5- To illustrate the robustness of the best fit procedure in this example we run a classical best fit procedure using the statistical package Stat: Fit. The procedure computes the MLE for the parameters of many distributions. Using these estimates it computes the p-value of the Kolmogorov-Smirnov goodness of fit test. Finally it ranks the distributions according to the p-values. The best fits according to this procedure were

Pearson type VI (42, 168, 9.3, 32.9) then

Extreme value(81.4,15.9)

Tiku and Akkaya (2004) showed that the MLE for censored sample is asymptotically equivalent to the robust MMLE (Modified Maximum Likelihood Estimator). So we remove the two outlying observations and repeat the procedure.The best fit distribution become

Pearson 3(46.9, 6, 7.26)

Pearson 5(14.1,19.6,1.43e+3)

Thus, we get the same best fit using the L-moments ratio diagram and the MLE for censored samples. In fact Pearson type III has been used extensively in hydrology frequency modeling, see Singh (1987) and Hosking and Wallis (1993).

References

1] Asquith, W. H., (2011) “lmomco package” .At http://cran.r-project-.org/web/packages/lmomco-/lmo-mco.pdf

2] Bajorunaite, R –Brazauskas, V. (2008). Method of trimmed moments for robust fitting of parametric failure time models. Metron, LXVI(3), 341-360.

3] Brazauskas, V., Jones, B. L., and Zitikisb, R. (2009). Robust fitting of claim severity distributions and the method of trimmed moments. Journal of Statistical Planning and Inference 139, 2028 -- 2043

4] Brys, G. , Hubert, M. and STruyf, A. (2004). A Robust Measure of Skewness. *J. Comput. and Graphical Stat., Vol. 13, pp 996–1017.*

5] Brys, G. , Hubert, M. and STruyf, A. (2008). Goodness- of-fit tests based on a robust measure of skewness. Comput. Statis. 27,429-442.

6] Dalen,J. (1987). Algebraic bounds on standardized sample moments. Statist. Prob. Lett. ,5,329-331.

7] D'Agostino, R. B., and Stephens, M.A., (1986)."Goodness of fit techniques". Marcel Decker, New York.

8] Elamir, E.A.H. and Seheult, A.H., (2003). "Trimmed L-moments". Computational Statistics and Data Analysis. 43, 299–314.

9] Elamir,E.A.H. and Seheult, A. H.,(2004). "Exact variance structure of sample L-moments" J. of statistical planning and inference, 124, 337-359.

10] Hampel, F. R. (1985). The breakdown points of the mean combined with some rejection rules. Technometrics 27, 95-107.

11] Hampel, F. R., Rousseeuw, P. J., Ronchetti, E. M., and Stahel, W. A. (1986). Robust Statistics. The Approach Based on Influence Function. Wiley, New York.

12] Heritier, S. and Ronchetti, E. M. (1994).Robust Bounded-Influence Tests in General Parametric Models. J. Amer. Statist. Assoc. 89, 897-904.

13] Heritier, S. and Victoria-Feser, M-P. (1997). “Practical Applications of Bounded-Influence Tests” Handbook of Statistics, Vol. 15, 76-100.

14] Hosking, J.R.M., (1990)." L-moments: analysis and estimation of distributions using linear combination of order statistics". J. Roy. Statist. Soc. B 52 (1), 105–124.

15] Hosking, J.R.M. (2006). "On the characterization of distributions by their L-moments" J. of statistical planning and inference,136, 193-198.

16] Hosking, J.R.M. (2007). "Some theory and practical uses of trimmed L-moments" J. of statistical planning and inference, 137, 3024-3039.

17] Hosking, J. R. and Wallis, J. R. (1993).“Regional Frequency Analysis: An Approach Based on L-moments”. Cambridge University Press. Cambridge.

18] Huber, P. J. and Ronchetti, E. M (2009). Robust Statistics, 2ed ed. Wiley, New York.

19] Hussien, Osama. A. (2011). the influence functions and breakdown points of the L-moments, and related moments estimators. The Egyptian Statistical Journal, ISSR

,Cairo University, vol. 55, No.1, pp

20] Hussien, Osama, (2008). “Robust Best Fit for Frequency Distributions”. Journal of the Egyptian Statistical Society,(24),1,80-113.

21] Hussien, Osama and Afifi, Nehal (2009). “Estimation and Goodness of Fit for the Generalized Lambda Distribution”. The Scientific Journal of the Faculty of Commerce, Tanat University.

22] Karvanen, J.(2006). "Estimation of quantile mixtures via L-moments and trimmed L-moments". Computational Statistics and Data Analysis, 51,947-959.

23] Krasker, W. S. and Welsch, R. E. (1982). Efficient bounded-influence regression estimation. J. Am. Statist. Assoc. 77, 595-604.

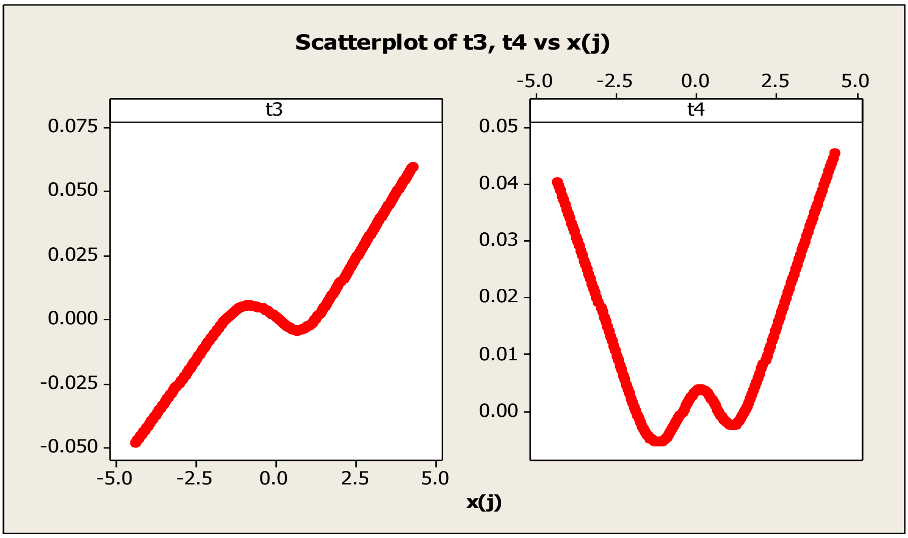
24] Michael, J. R. and Schucany, W. (1985). “The influence curve and goodness of fit”. J. American Statist. Assoc.(80) 678-682.

25] Ronchetti, E. (1997) “Robust inference by influence functions”. J. Statist. Plann. and Inference, 57, 59-72.

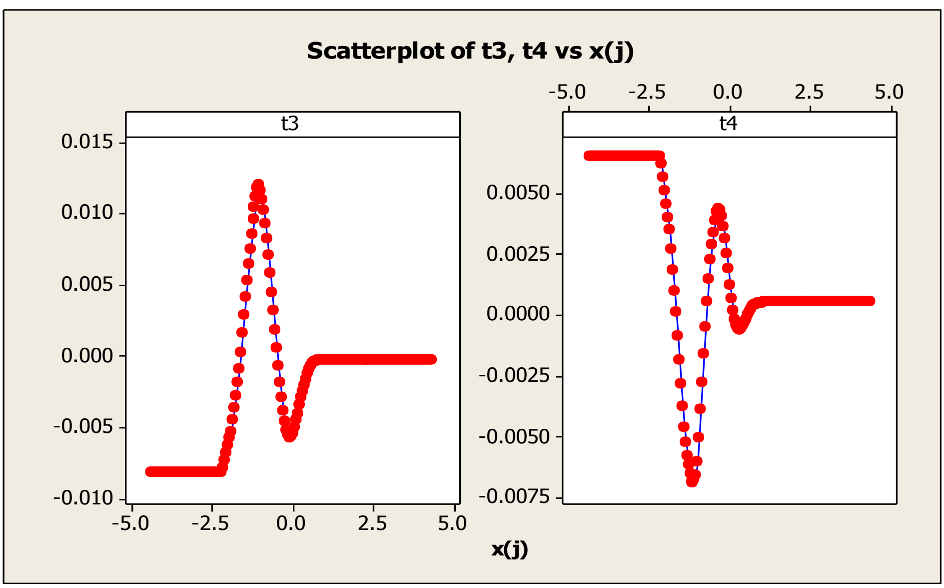
26] Ronchetti, E. and Victoria-Feser, M-P (1994). Robust Methods for Personal Income Distribution Models. The Canadian Journal of Statistics, Vol. 22, pp. 247-258

27] Atiem, I. A. and Harmancioglu, N.,(2006)"Assessment of regional floods using L-moments approach: The case of the river Nile". Water Resources Management 20:723-747.

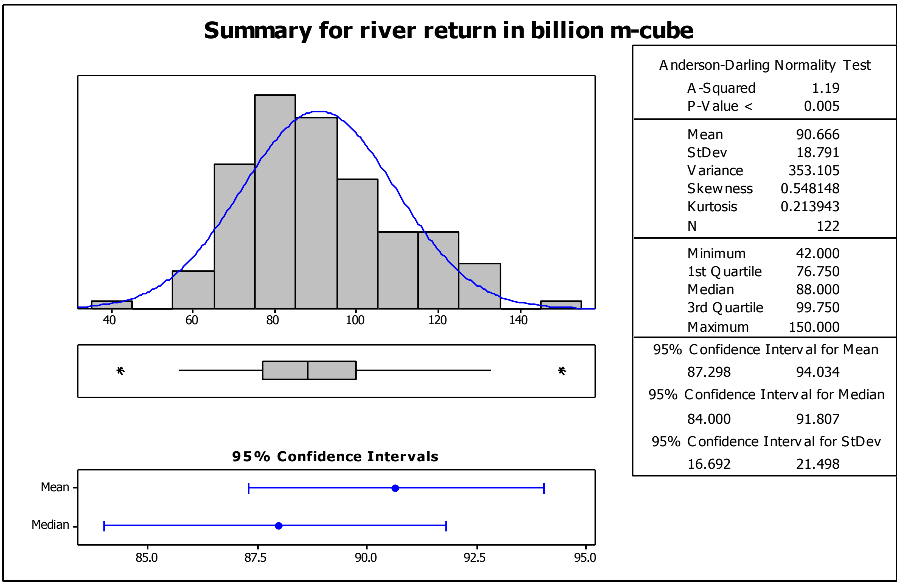
28] Tiku, M.L. Tan, W. Y. and Balakrishnan, N. (1986). “Robust Inference”.Mercel Dekker,New York.



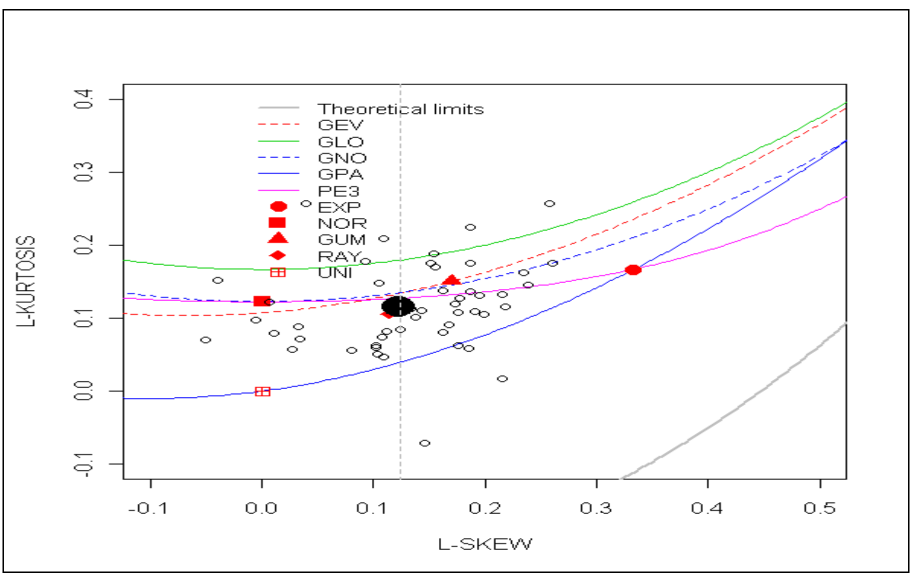
Graph 1 the influence function of L-skewness and L-kurtosis for the standard normal distribution



Graph 2 the influence function of TL-skewness andT L-kurtosis for the standard normal distribution



Graph (3) statistical summary of River Nile data



Graph (4) L-moment ratio diagram for the River Nile flood data. Sample (τ3 &τ4) marked as a large circle in the middle