J. R. Statist. Soc. B (1990)
52, *No.* 1, *pp.* 105–124

L-moments: Analysis and Estimation of Distributions using Linear Combinations of Order Statistics

By J. R. M. HOSKING†

IBM Research Division, Yorktown Heights, USA

[Received January 1989]

SUMMARY

L-moments are expectations of certain linear combinations of order statistics. They can be defined for any random variable whose mean exists and form the basis of a general theory which covers the summarization and description of theoretical probability distributions, the summarization and description of observed data samples, estimation of parameters and quantiles of probability distributions, and hypothesis tests for probability distributions. The theory involves such established procedures as the use of order statistics and Gini's mean difference statistic, and gives rise to some promising innovations such as the measures of skewness and kurtosis described in Section 2, and new methods of parameter estimation for several distributions. The theory of L-moments parallels the theory of (conventional) moments, as this list of applications might suggest. The main advantage of L-moments over conventional moments is that L-moments, being linear functions of the data, suffer less from the effects of sampling variability: L-moments are more robust than conventional moments to outliers in the data and enable more secure inferences to be made from small samples about an underlying probability distribution. L-moments sometimes yield more efficient parameter estimates than the maximum likelihood estimates.

Keywords: ESTIMATION; HYPOTHESIS TESTING; KURTOSIS; L-STATISTICS; MOMENTS; ORDER STATISTICS; SKEWNESS

1. INTRODUCTION

It is standard statistical practice to summarize a probability distribution or an observed data set by its moments or cumulants. It is also common, when fitting a parametric distribution to a data set, to estimate the parameters by equating the sample moments to those of the fitted distribution. Yet moment-based methods, although long established in statistics, are not always satisfactory. It is sometimes difficult to assess exactly what information about the shape of a distribution is conveyed by its moments of third and higher order; the numerical values of sample moments, particularly when the sample is small, can be very different from those of the probability distribution from which the sample was drawn; and the estimated parameters of distributions fitted by the method of moments are often markedly less accurate than those obtainable by other estimation procedures such as the method of maximum likelihood.

The alternative approach described here is based on quantities which we call L-moments. These are analogous to the conventional moments but can be estimated by linear combinations of order statistics, i.e. by L-statistics. L-moments have the

†*Address for correspondence*: IBM Research Division, T. J. Watson Research Center, PO Box 218, Yorktown Heights, NY 10598, USA.

 0035–9246/90/52105 $2.00

theoretical advantages over conventional moments of being able to characterize a wider range of distributions and, when estimated from a sample, of being more robust to the presence of outliers in the data. Experience also shows that, compared with conventional moments, L-moments are less subject to bias in estimation and approximate their asymptotic normal distribution more closely in finite samples. Parameter estimates obtained from L-moments are sometimes more accurate in small samples than even the maximum likelihood estimates.

Many statistical techniques are based on the use of linear combinations of order statistics: see David (1981) for examples. However, there has not heretofore been developed a unified approach to the use of order statistics for the statistical analysis of univariate probability distributions. We shall show that L-moments form the basis for such an approach, which covers the characterization of probability distributions, the summarization of observed data samples, the fitting of probability distributions to data and the testing of hypotheses about distributional form. Few of the theoretical results are new, being instead extensions of some scattered results and techniques described principally by Gini (1912), Sillitto (1951, 1969), Downton (1966), Chan (1967), Konheim (1971), Mallows (1973) and Greenwood *et al.* (1979). It is the gathering of these results into a unified whole, and the demonstration that L-moment methods perform competitively with the best available statistical techniques, which makes L-moments now worthy of the attention of statisticians.

2. L-MOMENTS OF PROBABILITY DISTRIBUTIONS

2.1. *Definitions and Basic Properties*

Let X be a real-valued random variable with cumulative distribution function $F(x)$ and quantile function $x(F)$, and let $X_{1:n} \leqslant X_{2:n} \leqslant \ldots \leqslant X_{n:n}$ be the order statistics of a random sample of size n drawn from the distribution of X. Define the *L-moments* of X to be the quantities

$$\lambda_r \equiv r^{-1} \sum_{k=0}^{r-1} (-1)^k \binom{r-1}{k} EX_{r-k:r}, \qquad r = 1, 2, \ldots. \tag{2.1}$$

The L in 'L-moments' emphasizes that λ_r is a *linear* function of the expected order statistics. Furthermore, as will be seen in Section 3, the natural estimator of λ_r based on an observed sample of data is a linear combination of the ordered data values, i.e. an L-statistic. The expectation of an order statistic may be written as

$$EX_{j:r} = \frac{r!}{(j-1)!(r-j)!} \int x\{F(x)\}^{j-1}\{1-F(x)\}^{r-j}\, \mathrm{d}F(x)$$

(David (1981), p. 33). Substituting this expression in definition (2.1), expanding the binomials in $F(x)$ and summing the coefficients of each power of $F(x)$ gives

$$\lambda_r = \int_0^1 x(F)\, P^*_{r-1}(F)\, \mathrm{d}F, \qquad r = 1, 2, \ldots, \tag{2.2}$$

where

$$P_r^*(F) = \sum_{k=0}^{r} p_{r,k}^* F^k \tag{2.3}$$

and

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

$P_r^*(F)$ is the rth shifted Legendre polynomial, related to the usual Legendre polynomials $P_r(u)$ by $P_r^*(u) = P_r(2u-1)$. Shifted Legendre polynomials are orthogonal on the interval (0, 1) with constant weight function (Lanczos (1957)—though his $P_r^*(\cdot)$ differs by a factor $(-1)^r$ from ours). The first few L-moments are

$$\begin{aligned}
\lambda_1 &= EX = \int_0^1 x(F)\, \mathrm{d}F, \\
\lambda_2 &= \tfrac{1}{2}E(X_{2:2} - X_{1:2}) = \int_0^1 x(F)\,(2F-1)\, \mathrm{d}F, \\
\lambda_3 &= \tfrac{1}{3}E(X_{3:3} - 2X_{2:3} + X_{1:3}) = \int_0^1 x(F)\,(6F^2 - 6F + 1)\, \mathrm{d}F, \\
\lambda_4 &= \tfrac{1}{4}E(X_{4:4} - 3X_{3:4} + 3X_{2:4} - X_{1:4}) = \int_0^1 x(F)\,(20F^3 - 30F^2 + 12F - 1)\, \mathrm{d}F.
\end{aligned} \tag{2.4}$$

The use of L-moments to describe probability distributions is justified by the following theorem.

Theorem 1.

(a) The L-moments λ_r, $r = 1, 2, \ldots$, of a real-valued random variable X exist if and only if X has finite mean.
(b) A distribution whose mean exists is characterized by its L-moments $\{\lambda_r: r = 1, 2, \ldots\}$.

Proof. A finite mean implies finite expectations of all order statistics (David (1981), p. 33), whence part (a) follows immediately. For part (b), let

$$\xi_r \equiv EX_{r:r} = r \int x\{F(x)\}^{r-1}\, \mathrm{d}F(x), \tag{2.5}$$

so that, from equation (2.2),

$$\begin{aligned}
\lambda_r &= \sum_{k=1}^{r} p_{r-1,k-1}^* k^{-1} \xi_k, \\
\xi_r &= \sum_{k=1}^{r} \frac{(2k-1)r!(r-1)!}{(r-k)!(r-1+k)!} \lambda_k.
\end{aligned} \tag{2.6}$$

Chan (1967) and Konheim (1971) proved that a distribution with finite mean is characterized by the set $\{\xi_r: r = 1, 2, \ldots\}$. By equations (2.6), a given set of λ_r

determines a unique set of ξ_r, so the characterization of a distribution in terms of the latter quantities extends to the former.

Thus a distribution may be specified by its L-moments even if some of its conventional moments do not exist. Furthermore, such a specification is always unique: this is of course not true of conventional moments.

As will shortly be shown, λ_2 is a measure of the scale or dispersion of the random variable X. It is often convenient to standardize the higher moments λ_r, $r \geqslant 3$, so that they are independent of the units of measurement of X. Define, therefore, the *L-moment ratios* of X to be the quantities

$$\tau_r \equiv \lambda_r/\lambda_2, \qquad r = 3, 4, \ldots .$$

It is also possible to define a function of L-moments which is analogous to the coefficient of variation (CV): this is the *L-CV*, $\tau \equiv \lambda_2/\lambda_1$. Bounds on the numerical values of the L-moment ratios and L-CV are given by the following theorem, proved in Hosking (1989).

Theorem 2. Let X be a non-degenerate random variable with finite mean. Then the L-moment ratios of X satisfy $|\tau_r| < 1$, $r \geqslant 3$. If in addition $X \geqslant 0$ almost surely, then τ, the L-CV of X, satisfies $0 < \tau < 1$.

We consider the boundedness of L-moment ratios to be an advantage. Intuitively, it is easier to interpret a measure such as τ_3, which is constrained to lie within the interval $(-1, 1)$, than the conventional skewness, which can take arbitrarily large values.

More stringent bounds on the τ_r can be found, via the ξ_r defined in equation (2.5). A complete specification of the possible values of the ξ_r, in terms of the positivity of the determinants of certain matrices whose elements are linear combinations of the ξ_r, is given by Mallows (1973), theorem 2(ii). From that theorem and equations (2.6) we can obtain the possible values of the L-moments of a distribution. In particular, for a non-degenerate distribution the constraints on ξ_1, ξ_2, ξ_3 and ξ_4 are that

$$\xi_2 - \xi_1 > 0, \qquad \xi_3 - \xi_2 > 0, \qquad -\xi_3 + 2\xi_2 - \xi_1 > 0,$$
$$(\xi_4 - \xi_3)(\xi_2 - \xi_1) - (\xi_3 - \xi_2)^2 \geqslant 0, \qquad -\xi_4 + 2\xi_3 - \xi_2 > 0,$$

and so the constraints on λ_1, λ_2, τ_3 and τ_4 are that

$$0 < \lambda_2, \qquad -1 < \tau_3 < 1, \qquad \tfrac{1}{4}(5\tau_3^2 - 1) \leqslant \tau_4 < 1. \tag{2.7}$$

2.2. *Probability Weighted Moments*

Greenwood *et al.* (1979) defined probability weighted moments (PWMs) to be the quantities

$$M_{p,r,s} = E[X^p\{F(X)\}^r\{1 - F(X)\}^s],$$

and they and others (Landwehr *et al.*, 1979a, b; Wallis, 1980; Greis and Wood, 1981; Hosking *et al.*, 1985; Hosking and Wallis, 1987a) developed statistical inference procedures using the PWMs $M_{1,0,s}$ and $M_{1,r,0}$. These PWMs can be expressed as linear combinations of L-moments, so procedures based on PWMs and on L-moments are equivalent. L-moments are more convenient, however, because they are more directly interpretable as measures of the scale and shape of probability distributions.

2.3. *Summarizing a Probability Distribution*

The L-moments $\lambda_1, \ldots, \lambda_r$ and the L-moment ratios $\tau_3, \ldots, \tau_r$ are useful quantities for summarizing a distribution. The L-moments are in some ways analogous to the (conventional) central moments and the L-moment ratios are analogous to moment ratios. In particular, λ_1, λ_2, τ_3 and τ_4 may be regarded as measures of location, scale, skewness and kurtosis respectively.

To see this consider equations (2.4), the definition of the λ_r as expectations of linear combinations of order statistics. Clearly λ_1, the mean, is a measure of location. To interpret λ_2, consider the typical configuration of a sample of size 2: if the two values tend to be close together, as in Fig. 1(a), then λ_2 will be smaller than if they are far apart, as in Fig. 1(b). Thus λ_2 can be thought of as measuring the scale or dispersion of the distribution. Samples of size 3 are relevant to λ_3, which is the central second difference of the median of such a sample. Samples like that of Fig. 1(c), which yields a positive central second difference, tend to arise from positively skew distributions; Fig. 1(d) is more typical of distributions with negative skewness. Symmetric distributions have $\lambda_3 = 0$. Thus λ_3 may be thought of as measuring skewness, although not independently of scale. Similarly, (e) and (f) of Fig. 1 illustrate samples of size 4. Configuration (e), typical of a heavy-tailed or sharply peaked distribution, has a large positive central third difference, while configuration (f), more typical of a flat or even U-shaped distribution, has a negative central third difference. Thus λ_4, itself the central third difference of the expected order statistics of a sample of size 4, measures the same aspects of a distribution as does the fourth central (conventional) moment. The L-moment ratios τ_3 and τ_4 are dimensionless analogues of λ_3 and λ_4 respectively and are therefore plausible measures of skewness and kurtosis.

An alternative justification of these interpretations of L-moments may be based on the work of Oja (1981). Extending work of Bickel and Lehmann (1975, 1976) and van Zwet (1964), Oja defined intuitively reasonable criteria for one probability distribution on the real line to be located further to the right (more dispersed, more skew, more kurtotic) than another. A real-valued functional of a distribution that preserves the partial ordering of distributions implied by these criteria may then reasonably be called a 'measure of location' (dispersion, skewness, kurtosis). It follows immediately from Oja's work that λ_1 and λ_2, in Oja's notation $\mu_1(F)$ and $\frac{1}{2}\sigma_1(F)$, are measures of location and scale respectively. Hosking (1989) shows that τ_3 and τ_4 are, by Oja's criteria, measures of skewness and kurtosis respectively.

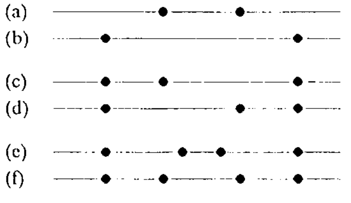


Fig. 1. Configurations of samples of sizes 2, 3 and 4

2.4. *L-skewness and L-kurtosis*

Section 2.3 implies that the main features of a probability distribution should be well summarized by the following four measures: the mean or *L-location*, λ_1; the *L-scale*, λ_2; the *L-skewness*, τ_3; the *L-kurtosis*, τ_4. We now consider these measures, particularly τ_3 and τ_4, in more detail.

The L-moment measure of location is the mean, λ_1. This is a well-established and familiar quantity which needs no further description or justification here.

The L-scale λ_2 is also long established in statistics, for it is, apart from a scalar multiple, the expectation of Gini's mean difference statistic. To compare λ_2 with the more familiar scale measure σ, the standard deviation, write

$$\lambda_2 = \tfrac{1}{2}E(X_{2:2} - X_{1:2}), \qquad \sigma^2 = \tfrac{1}{2}E(X_{2:2} - X_{1:2})^2.$$

Both quantities measure the difference between two randomly drawn elements of a distribution, but σ^2 gives relatively more weight to the largest differences.

The L-skewness τ_3 is a dimensionless analogue of λ_3. By theorem 2, τ_3 takes values between -1 and $+1$. These bounds are the best possible: they are approached arbitrarily closely as $p \to 0$ or $p \to 1$ in the Bernoulli random variable X_p with $P[X_p = 0] = p$, $P[X_p = 1] = 1 - p$. Symmetric distributions have $\tau_3 = 0$. Apart from brief mentions by Sillitto (1951) and Kaigh and Driscoll (1987), τ_3 has not appeared previously in the statistical literature. However, using Sillitto (1951), equation (9), to write

$$\tau_3 = \frac{EX_{3:3} - 2EX_{2:3} + EX_{1:3}}{EX_{3:3} - EX_{1:3}}$$

shows that τ_3 is similar in form to a measure of skewness used by Bowley (1937):

$$B \equiv \frac{Q_3 - 2Q_2 + Q_1}{Q_3 - Q_1}$$

where $Q_r \equiv x(r/4)$, $r = 1, 2, 3$, are the quartiles of X. Skewness measures similar to B but based on quantiles other than the quartiles have been used by Hinkley (1975) and Groeneveld and Meeden (1984). As a measure of skewness, B can be criticized for being insensitive to the distribution of X any further into the tails than the quartiles. In contrast, the conventional moment-based measure of skewness,

$$\gamma \equiv E(X - EX)^3/\{E(X - EX)^2\}^{3/2},$$

is so sensitive to the extreme tails of the distribution that it is difficult to estimate accurately in practice when the distribution is markedly skew. We believe that the skewness measure τ_3 steers an advantageous middle course between these extremes.

It is interesting to compare the skewness measures τ_3 and γ for various distributions: the comparison is made graphically in Fig. 2. For symmetric distributions both τ_3 and γ are zero, and many near symmetric distributions have $\gamma \approx 6\tau_3$, but in general there is no simple relationship between γ and τ_3. Both γ and τ_3 may yield a large positive skewness either when a distribution has a heavy right tail or when a continuous distribution is reverse J shaped, i.e. has a finite lower bound near which $f(x) \to \infty$. The former case tends to yield particularly high values of γ relative to τ_3, because γ is more sensitive to the extreme tail weight of the distribution. Indeed for some heavy-tailed distributions γ approaches infinity while τ_3 has still quite a modest

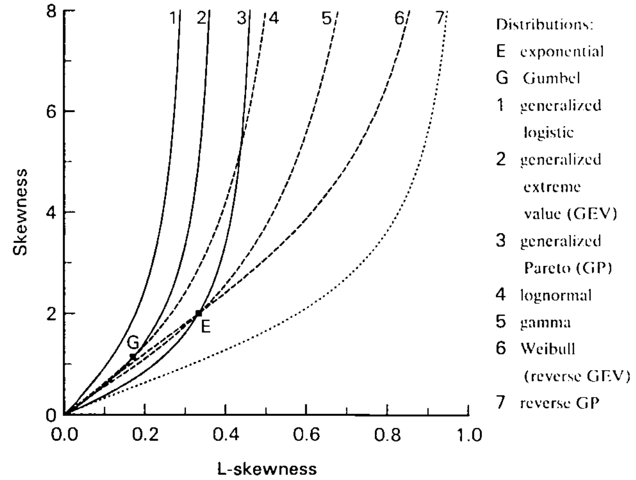


Fig. 2. Comparison of skewness and L-skewness: ——, ---, ····, distributions whose asymptotic distributions of maxima are of extreme value types II, I and III respectively, i.e., roughly, distributions with power law upper tails, exponential upper tails and finite upper limits respectively

value: for example, 0.25 for the generalized logistic distribution.

Kurtosis, as measured by the moment ratio

$$\kappa \equiv E(X - EX)^4/\{E(X - EX)^2\}^2,$$

has no unique interpretation. It can be thought of as the 'peakedness' of a distribution, or as 'tail weight', but only for fairly closely defined families of symmetric unimodal distributions do these interpretations have any demonstrable validity (see Balanda and MacGillivray (1988), and references therein). L-kurtosis, τ_4, is equally difficult to interpret uniquely and is best thought of as a measure similar to κ but giving less weight to the extreme tails of the distribution. To illustrate the greater dependence of κ on the extreme tails of a distribution, consider a lambda distribution (Tukey, 1960) with quantile function $x(F) = F^\lambda - (1 - F)^\lambda$ and $\lambda = -0.1466$. If the distribution is truncated at its 0.001 and 0.999 quantiles, its kurtosis κ falls from 10.00 to 5.48, but its L-kurtosis τ_4 falls only from 0.224 to 0.204.

Table 1 gives the first four L-moments and L-moment ratios for some common distributions. Values of τ_3 and τ_4 can be plotted to yield an L-moment ratio diagram, exemplified by Fig. 3. The 'bound for all distributions' is the last inequality of equations (2.7). The uniform distribution has $\tau_r = 0$ for all $r \geqslant 3$ and thus plays a central role in L-moment theory akin to that of the normal distribution in cumulant theory.

Higher L-moments may be viewed similarly to higher conventional moments. For example τ_5 could be interpreted as a measure of tendency to bimodality, while the τ_r of odd order are generalized skewness measures in so far as symmetric distributions have $\tau_{2r+1} = 0$ for all $r \geqslant 1$.

2.5. *Approximating a Quantile Function*

Sillitto (1969) derived L-moments, without so naming them, as coefficients in the approximation of a quantile function by polynomials. As a matter of taste, we prefer

TABLE 1
L-moments of some common distributions†

Distribution	*F(x) or x(F)*	*L-moments*
Uniform	$x=\alpha+(\beta-\alpha)F$	$\lambda_1=\frac{1}{2}(\alpha+\beta)$, $\lambda_2=\frac{1}{6}(\beta-\alpha)$, $\tau_3=0$, $\tau_4=0$
Exponential	$x=\xi-\alpha\log(1-F)$	$\lambda_1=\xi+\alpha$, $\lambda_2=\frac{1}{2}\alpha$, $\tau_3=\frac{1}{3}$, $\tau_4=\frac{1}{6}$
Gumbel	$x=\xi-\alpha\log(-\log F)$	$\lambda_1=\xi+\gamma\alpha$, $\lambda_2=\alpha\log 2$, $\tau_3=0.1699$, $\tau_4=0.1504$
Logistic	$x=\xi+\alpha\log\{F/(1-F)\}$	$\lambda_1=\xi$, $\lambda_2=\alpha$, $\tau_3=0$, $\tau_4=\frac{1}{6}$
Normal	$F=\Phi\left(\frac{x-\mu}{\sigma}\right)$	$\lambda_1=\mu$, $\lambda_2=\pi^{-1}\sigma$, $\tau_3=0$, $\tau_4=30\pi^{-1}\tan^{-1}\sqrt{2}-9=0.1226$
Generalized Pareto	$x=\xi+\alpha\{1-(1-F)^k\}/k$	$\lambda_1=\xi+\alpha/(1+k)$, $\lambda_2=\alpha/(1+k)(2+k)$, $\tau_3=(1-k)/(3+k)$, $\tau_4=(1-k)(2-k)/(3+k)(4+k)$
Generalized extreme value	$x=\xi+\alpha\{1-(-\log F)^k\}/k$	$\lambda_1=\xi+\alpha\{1-\Gamma(1+k)\}/k$, $\lambda_2=\alpha(1-2^{-k})\Gamma(1+k)/k$, $\tau_3=2(1-3^{-k})/(1-2^{-k})-3$, $\tau_4=(1-6.2^{-k}+10.3^{-k}-5.4^{-k})/(1-2^{-k})$
Generalized logistic	$x=\xi+\alpha[1-\{(1-F)/F\}^k]/k$	$\lambda_1=\xi+\alpha\{1-\Gamma(1+k)\,\Gamma(1-k)\}/k$, $\lambda_2=\alpha\Gamma(1+k)\,\Gamma(1-k)$, $\tau_3=-k$, $\tau_4=(1+5k^2)/6$
Log-normal	$F=\Phi\left(\frac{\log(x-\xi)-\mu}{\sigma}\right)$	$\lambda_1=\xi+\exp(\mu+\sigma^2/2)$, $\lambda_2=\exp(\mu+\sigma^2/2)\,\mathrm{erf}(\sigma/2)$, $\tau_3=6\pi^{-1/2}\int_0^{\sigma/2}\mathrm{erf}(x/\sqrt{3})\exp(-x^2)\mathrm{d}x/\mathrm{erf}(\sigma/2)$
Gamma	$F=\beta^{-\alpha}\int_0^x t^{\alpha-1}\exp(-t/\beta)\mathrm{d}t/\Gamma(\alpha)$	$\lambda_1=\alpha\beta$, $\lambda_2=\pi^{-1/2}\beta\Gamma(\alpha+\frac{1}{2})/\Gamma(\alpha)$, $\tau_3=6I_{1/3}(\alpha, 2\alpha)-3$

†γ is Euler's constant; Φ is the standard normal distribution function; $I_x(p, q)$ is the incomplete beta function. Expressions for τ_4 for the gamma and log-normal distributions are given by Hosking (1986).

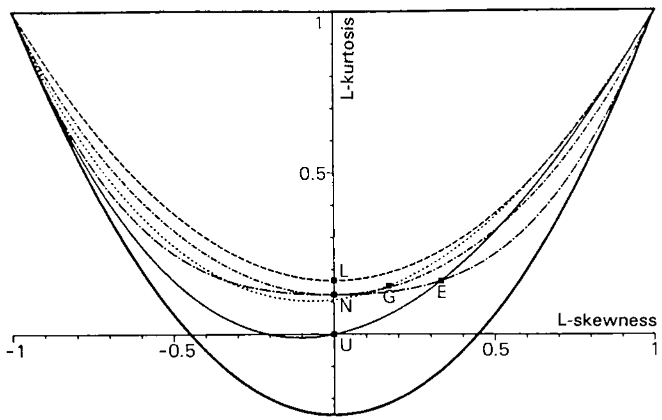
to regard equation (2.1) as the fundamental definition; the approximation to the quantile function then becomes an inversion theorem, expressing the quantile function in terms of the L-moments.

Theorem 3 (Sillitto, 1969). Let X be a real-valued continuous random variable with finite variance, quantile function $x(F)$ and L-moments λ_r, $r \geqslant 1$. Then the representation

$$x(F) = \sum_{r=1}^{\infty} (2r-1)\lambda_r P^*_{r-1}(F), \qquad 0 < F < 1,$$

is convergent in mean square.

The representation for $x(F)$ given by the inversion theorem is of limited practical utility. The approximation to $x(F)$ using a finite number of L-moments can be poor in the tails of the distribution, particularly if the distribution has a heavy tail; not uncommonly there are some intervals of F in which the approximation to $x(F)$ is not monotonic increasing. Similar problems arise with the Cornish–Fisher expansion of $x(F)$ in terms of the cumulants of X. Just as the Cornish–Fisher expansion is most useful for near normal distributions, we would expect the approximation of $x(F)$ by L-moments to be most accurate when the distribution of X is close to uniform.



Distributions: U uniform, E exponential, G Gumbel, L logistic, N Normal,
— generalized Pareto, ······ generalized extreme value, ---- generalized logistic,
-·-·- lognormal, —·— gamma, — bound for all distributions

Fig. 3. L-moment ratios of some common distributions

3. ESTIMATION OF L-MOMENTS

3.1. *Unbiased Estimators*

In practice, L-moments must usually be estimated from a random sample drawn from an unknown distribution. Because λ_r is a function of the expected order statistics of a sample of size r, it is natural to estimate it by a U-statistic, i.e. the corresponding function of the sample order statistics averaged over all subsamples of size r which can be constructed from the observed sample of size n. Let $x_1, x_2, \ldots, x_n$ be the sample and $x_{1:n} \leqslant x_{2:n} \leqslant \ldots \leqslant x_{n:n}$ the ordered sample, and define the rth *sample L-moment* to be

$$l_r = \binom{n}{r}^{-1} \sum_{1 \leqslant i_1 < i_2 < \ldots < i_r \leqslant n} \sum \ldots \sum r^{-1} \sum_{k=0}^{r-1} (-1)^k \binom{r-1}{k} x_{i_{r-k}:n}, \qquad r = 1, 2, \ldots, n; \tag{3.1}$$

in particular

$$l_1 = n^{-1} \sum_i x_i,$$

$$l_2 = \frac{1}{2}\binom{n}{2}^{-1} \sum_{i>j}\sum (x_{i:n} - x_{j:n}),$$

$$l_3 = \frac{1}{3}\binom{n}{3}^{-1} \sum_{i>j>k}\sum\sum (x_{i:n} - 2x_{j:n} + x_{k:n}),$$

$$l_4 = \frac{1}{4}\binom{n}{4}^{-1} \sum_{i>j>k>l}\sum\sum\sum (x_{i:n} - 3x_{j:n} + 3x_{k:n} - x_{l:n}).$$

U-statistics were introduced by Hoeffding (1948) and are widely used in non-parametric statistics (see, for example, Fraser (1957) and Randles and Wolfe (1979)). Their properties of unbiasedness, asymptotic normality and some modest resistance to the influence of outliers make them particularly attractive for statistical inference.

When calculating l_r it is not necessary to iterate over all subsamples of size r; the statistic can be expressed explicitly as a linear combination of order statistics of a sample of size n, as in Blom (1980). Considering the U-statistic estimator of $EX_{r:r}$, and counting the occurrences in it of each $x_{i:n}$, shows that the estimator may be written as rb_{r-1}, where, as in Hosking *et al.* (1985),

$$b_r = n^{-1} \sum_{i=1}^{n} \frac{(i-1)(i-2) \ldots (i-r)}{(n-1)(n-2) \ldots (n-r)} x_{i:n};$$

thus by equation (2.5)

$$l_r = \sum_{k=0}^{r-1} p^*_{r-1,k} b_k.$$

Sample L-moments may be used similarly to (conventional) sample moments: they summarize the basic properties—location, scale, skewness, kurtosis—of a data set, they estimate the corresponding properties of the probability distribution from which the data were sampled and they may be used to estimate the parameters of the underlying distribution. In these applications L-moments are frequently preferable to conventional moments: being linear functions of the data, they are less sensitive than are conventional moments to sampling variability or measurement errors in the extreme data values, and may therefore be expected to yield more accurate and robust estimates of the characteristics or parameters of an underlying probability distribution.

Under a linear transformation of the data, the sample L-moments are transformed isomorphically with the corresponding population L-moments. If $x_i \to Ax_i + B$, $i = 1, \ldots, n$, then $l_1 \to Al_1 + B$ and $l_r \to (\operatorname{sign} A)^r A l_r$, $r \geqslant 2$.

Sample L-moments have been used previously in statistics, although not as part of a unified theory. The statistic l_1 is the sample mean. The sample L-scale, l_2, is a scalar multiple of Gini's mean difference statistic

$$G = \binom{n}{2}^{-1} \sum\sum_{i>j} (x_{i:n} - x_{j:n})$$

which has been used in statistics since at least as far back as von Andrae (1872) and Gini (1912). The statistic $\frac{1}{2}\pi^{1/2}G$ is a 98% efficient estimator of the scale parameter of a normal distribution (Downton, 1966; David, 1968). G is also related to the 'total time on test' statistic for testing exponentiality (Gail and Gastwirth, 1978). The statistic l_3/s, where s is the sample standard deviation, has been used to test for normality by Locke and Spurrier (1976). However, it is more logical to base such a test on $t_3 = l_3/l_2$, as in Section 4.2 later.

3.2. *Plotting-position Estimators*

A *plotting position* is a distribution-free estimator of $F(x_{i:n})$. Reasonable choices for plotting positions include $p_{i:n} \equiv (i+\gamma)/(n+\delta)$ for $\delta > \gamma > -1$. Plotting positions are used in the graphical display of statistical data (Cunnane, 1978; Harter, 1984) but also provide a means of estimating quantities of the form $\int x(F)\ \eta(F)\ \mathrm{d}F$ where $x(F)$ is the quantile function of a distribution and η is a function of F alone. Equation (2.2) shows that λ_r is of this form, so it can be estimated by

$$\tilde{\lambda}_r \equiv \sum_{i=1}^{n} P^*_{r-1}(p_{i:n})x_{i:n}.$$

In general $\tilde{\lambda}_r$ is not an unbiased estimator of λ_r, but it is consistent. Indeed $\tilde{\lambda}_r$ and l_r are asymptotically equivalent, the difference between them being of stochastic order n^{-1}.

There is no theoretical reason for preferring plotting-position estimators to the unbiased estimators, but practical experience shows that plotting-position estimators sometimes yield better estimates of parameters and quantiles when a distribution is fitted to data. In particular the choice $p_{i:n} = (i-0.35)/n$ gives good results for the generalized Pareto (Hosking and Wallis, 1987a), generalized extreme value (GEV) (Hosking *et al.*, 1985) and Wakeby (Landwehr *et al.*, 1979b) distributions.

3.3. *Estimation of L-moment Ratios*

The L-moment ratio $\tau_r = \lambda_r/\lambda_2$ is naturally estimated by $t_r \equiv l_r/l_2$. The plotting-position estimator $\tilde{\tau}_r \equiv \tilde{\lambda}_r/\tilde{\lambda}_2$ is asymptotically equivalent to t_r and will not be considered separately. Analogously to our previous terminology, t_r is called the rth *sample L-moment ratio*, t_3 is the *sample L-skewness* and t_4 is the *sample L-kurtosis.*

The sample L-moment ratios t_3 and t_4 may be used to measure the skewness and kurtosis of an observed data set, as is commonly done with the conventional sample moment ratios g (skewness) and k (kurtosis). A disadvantage of the conventional moment ratios, noted by Kirby (1974) and Dalén (1987), is that when calculated from finite samples they are bounded and cannot attain the full range of values available to the population skewness and kurtosis. For example the skewness g is bounded by

$$|g| \leqslant (n-2)/(n-1)^{1/2}$$

for a sample of size n, and for many moderately to highly skew distributions it is unusual for g to take a value anywhere near the population skewness γ (Wallis *et al.*, 1974). In contrast, it can be shown (Hosking, 1986) that the sample L-moment ratios (t_3, t_4) calculated from a sample of size $n \geqslant 4$ can take any of the feasible values of the population L-moment ratios (τ_3, τ_4).

3.4. *Sampling Distributions of L-moments*

Exact sampling distributions of L-moments are difficult to obtain. The most practically useful results come from asymptotic distribution theory. Asymptotic theory for linear combinations of order statistics, developed by Chernoff *et al.* (1967), Moore (1968) and Stigler (1974) among others, can be immediately applied to sample L-moments and L-moment ratios. The main result is the following theorem.

Theorem 3. Let X be a real-valued random variable with cumulative distribution function F, L-moments λ_r and finite variance. Let l_r, $r = 1, 2, \ldots, m$, be sample L-moments calculated from a random sample of size n drawn from the distribution of X. Let $\tau_r = \lambda_r/\lambda_2$ and $t_r = l_r/l_2$, $r = 3, 4, \ldots, m$. Then as $n \rightarrow \infty$:

(a) $n^{1/2}(l_r-\lambda_r)$, $r = 1, 2, \ldots, m$, converge in distribution to the multivariate normal distribution $N(0, \Lambda)$, where the elements Λ_{rs} $(r, s = 1, 2, \ldots, m)$ of Λ are given by

$$\Lambda_{rs} = \iint_{x<y} \{P^*_{r-1}(F(x))P^*_{s-1}(F(y)) + P^*_{s-1}(F(x))P^*_{r-1}(F(y))\}\, F(x)\{1-F(y)\}\, \mathrm{d}x\, \mathrm{d}y, \tag{3.2}$$

$P^*_r(x)$ being the rth shifted Legendre polynomial defined in equation (2.3);

(b) the vector

$$n^{1/2}[(l_1-\lambda_1)\ (l_2-\lambda_2)\ (t_3-\tau_3)\ (t_4-\tau_4) \ldots (t_m-\tau_m)]^{\mathrm{T}}$$

converges in distribution to the multivariate normal distribution $N(0, T)$ where the elements T_{rs} $(r, s = 1, 2, \ldots, m)$ of T are given by

$$T_{rs} = \begin{cases} \Lambda_{rs} & \text{if } r \leqslant 2, s \leqslant 2, \\ (\Lambda_{rs}-\tau_r\Lambda_{2s})/\lambda_2 & \text{if } r \geqslant 3, s \leqslant 2, \\ (\Lambda_{rs}-\tau_r\Lambda_{2s}-\tau_s\Lambda_{2r}+\tau_r\tau_s\Lambda_{22})/\lambda_2^2 & \text{if } r \geqslant 3, s \geqslant 3. \end{cases}$$

Proof.

(a) Theorem 6 of Stigler (1974) gives the asymptotic normal distribution of linear combinations of order statistics and can be applied immediately to any finite linear combination of the statistics $\tilde{\lambda}_r$ based on the plotting position $i/(n+1)$. Thus these $\tilde{\lambda}_r$ have an asymptotic normal distribution with covariance matrix, as evaluated using Stigler's theorem, given by Λ as defined above. By expressing l_r as a sum of the $\tilde{\lambda}_s$, $s = 1, \ldots, r$, it is straightforward to show that $l_r - \tilde{\lambda}_r = O_{\mathrm{p}}(n^{-1})$; thus the asymptotic distributions of the l_r and of the $\tilde{\lambda}_r$ are identical.

(b) The result follows from standard theory for functions of asymptotically normally distributed vectors (Serfling (1980), section 3.3).

When the random variable X is continuous, with a differentiable quantile function $x(F)$, Λ_{rs} may be written as

$$\Lambda_{rs} = \iint_{0<u<v<1} \{P^*_{r-1}(u)\, P^*_{s-1}(v) + P^*_{s-1}(u)\, P^*_{r-1}(v)\}\, u(1-v)\, x'(u)\, x'(v)\, \mathrm{d}u\, \mathrm{d}v. \tag{3.3}$$

This is often the most convenient form for evaluating these quantities for specific distributions. Hosking (1986) gives several examples.

Asymptotic biases of L-moment ratios can be calculated by standard Taylor series expansion methods (e.g. Rao (1973)). For most distributions these biases are negligible for sample sizes of 20 or more. For example, for the normal distribution ($\tau_4 = 0.1226$) the asymptotic bias of t_4 is $0.03n^{-1}$, and for the logistic distribution ($\tau_4 = 0.1667$) this bias is $-0.04n^{-1}$.

Asymptotic theory is useful only in so far as it yields accurate approximations to finite sample distributions. Fig. 4 shows the distributions of skewness and L-skewness for samples of size 50 drawn from the Gumbel distribution, together with the

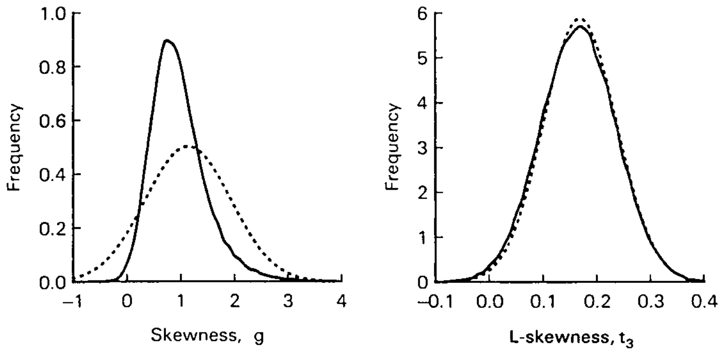


Fig. 4. ——, distributions of sample skewness g and sample L-skewness t_3 for samples of size 50 from the Gumbel distribution; ---, approximations to these distributions based on asymptotic theory (results for sample size 50 were obtained by Monte Carlo simulation using 100000 simulated samples)

asymptotic normal approximations to these distributions. The normal approximation to the distribution of the sample L-skewness t_3 is very good; for the conventional moment statistic g it is unusably poor.

Of course, the true devotee of L-moments would prefer to describe the limiting normal distributions of sample L-moments in terms of their L-moments rather than their means and variances. Since a random variable distributed as $N(\mu, \sigma^2)$ has L-moments $\lambda_1 = \mu$ and $\lambda_2 = \pi^{-1/2}\sigma$ this is easily done. An L-moment analogue of covariance, however, is not so easy to define.

3.5. *Identification of Distributions*

An important application of summary statistics calculated from an observed random sample is to identify the distribution from which the sample was drawn. This is much more easily achieved, particularly for skew distributions, by using L-moments rather than conventional moments.

As an example, 50 random samples of size 100 were simulated from each of three distributions: a GEV distribution with skewness 3 and two Weibull distributions, one with the same skewness and one with the same L-skewness as this GEV distribution. The distributions are illustrated in Fig. 5. Moments and L-moments of the generated samples are shown in Fig. 6. The sample conventional moments from the three distributions all lie close to a single line on the graph and overlap each other; they offer little hope of identifying the population distribution. In contrast, the sample L-moments plot as fairly well separated groups and permit a high probability of discrimination between the three distributions.

A practical example of distinguishing between distributions has been discussed by Hosking and Wallis (1987b). Analysis by the State of California Department of Water Resources (1981) of annual maximum hourly rainfall data at 689 rain-gauges in California suggested that a gamma distribution was appropriate for the data because the average values of sample skewness and kurtosis were consistent with the relationship $\kappa = 3 + 3\gamma^2/2$ between the population skewness and kurtosis of gamma distributions. This inference is valid only if sample moments are accurate estimators of population moments, which for the California data is dubious because the sample

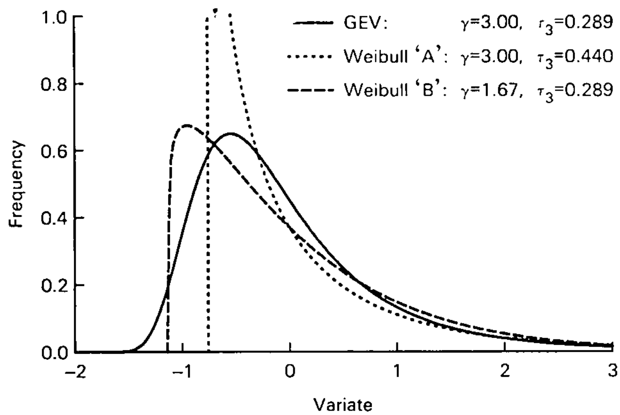


Fig. 5. Probability density functions of three probability distributions: each distribution has zero mean and unit variance

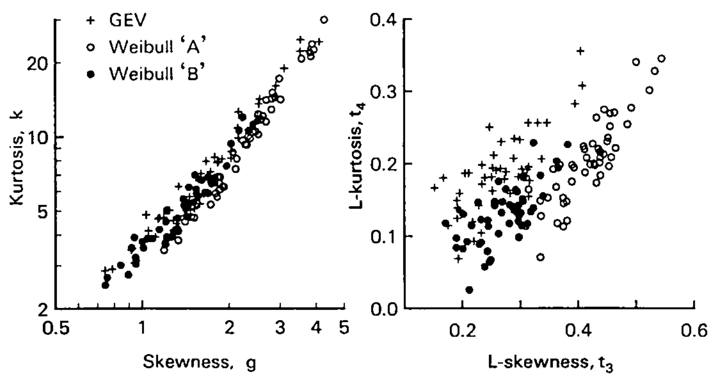


Fig. 6. Sample skewness versus sample kurtosis and sample L-skewness versus sample L-kurtosis, for examples of size 100 simulated from the three distributions of Fig. 5

sizes are small—only 12 gauges have records for as long as 50 years. L-moments tell a different story. The sample t_3 and t_4 values for the 68 sites with at least 20 years of records in the Central Valley of California are shown in Fig. 7(a). The data are on average closer to the population L-moments of a GEV distribution rather than a gamma distribution. Furthermore the spread of the data about the GEV line is consistent with what we might expect from data sampled from independent GEV distributions each with $\tau_3 = 0.24$ (the average of the 68 values of t_3) and the same record lengths as the actual rainfall records—see Fig. 7(b). Sample L-moments of data simulated from a gamma distribution are shown in Fig. 7(c), which has an appearance rather different from Figs 7(a) and 7(b): there are many fewer points above the GEV line and many more below the gamma distribution line. Thus the Central Valley hourly rainfall data may be well described by a GEV distribution but seem most unlikely to follow a gamma distribution.

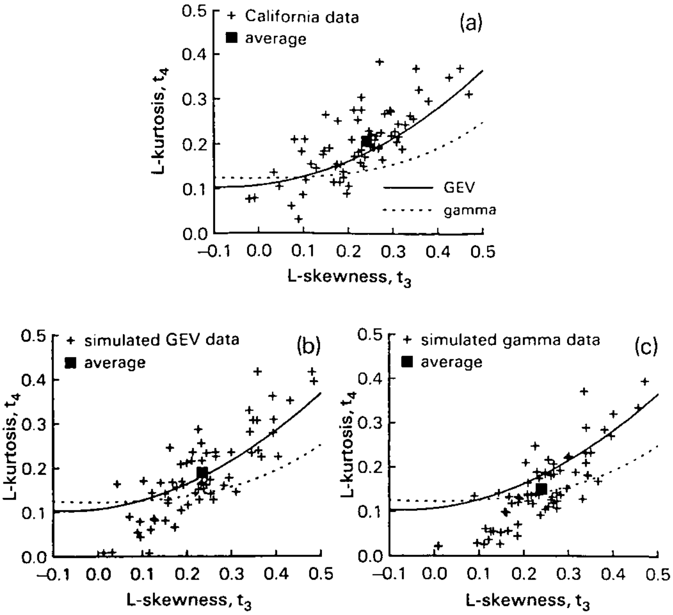


Fig. 7. (a) Values of t_3 and t_4 for rainfall data from 68 sites in the Central Valley of California, together with the theoretical relationship between τ_3 and τ_4 for GEV and gamma distributions; (b) values of t_3 and t_4 for 68 samples, with sample sizes the same as for the Central Valley data, simulated from a GEV distribution with $\tau_3=0.24$; (c) values of t_3 and t_4 for 68 samples, with sample sizes the same as for the Central Valley data, simulated from a gamma distribution with $\tau_3=0.24$

4. PARAMETER ESTIMATION AND HYPOTHESIS TESTING USING L-MOMENTS

4.1. *Parameter Estimation*

A common problem in statistics is the estimation, from a random sample of size n, of a probability distribution whose specification involves a finite number p of unknown parameters. Analogously to the usual method of moments, the 'method of L-moments' obtains parameter estimates by equating the first p sample L-moments to the corresponding population quantities. Examples of parameter estimators derived using this method are given in Table 2.

Exact distributions of parameter estimators obtained by the method of L-moments are in general difficult to derive. Asymptotic distributions can be found by treating the estimators as functions of sample L-moments and applying Taylor series methods (Serfling (1980), p. 122). Hosking (1986) gives several examples of such results. For most standard distributions, this approach can be used to show that L-moment estimators of parameters and quantiles are asymptotically normally distributed and to find standard errors and confidence intervals. In applications we have found that asymptotic approximations are usually reliable for samples of size 50 or more: see, for example, Hosking *et al.* (1985) and Hosking and Wallis (1987a).

It is of interest to compare the method of L-moments with the asymptotically optimal method of maximum likelihood. The method of L-moments is usually computationally more tractable than the method of maximum likelihood and needs less frequent recourse to iterative procedures. The asymptotic standard errors of

TABLE 2

Parameter estimation via L-moments for some common distributions†

Distribution	*Estimators*
Exponential	(ξ known) $\hat{\alpha}=l_1$
Gumbel	$\hat{\alpha}=l_2/\log 2$, $\hat{\xi}=l_1-\gamma\hat{\alpha}$
Logistic	$\hat{\alpha}=l_2$, $\hat{\xi}=l_1$
Normal	$\hat{\sigma}=\pi^{1/2}l_2$, $\hat{\mu}=l_1$
Generalized Pareto	(ξ known) $\hat{k}=l_1/l_2-2$, $\hat{\alpha}=(1+\hat{k})l_1$
Generalized extreme value	$z=2/(3+t_3)-\log 2/\log 3$, $\hat{k}\approx 7.8590z+2.9554z^2$, $\hat{\alpha}=l_2\hat{k}/(1-2^{-\hat{k}})\Gamma(1+\hat{k})$, $\hat{\xi}=l_1+\hat{\alpha}\{\Gamma(1+\hat{k})-1\}/\hat{k}$
Generalized logistic	$\hat{k}=-t_3$, $\hat{\alpha}=l_2/\Gamma(1+\hat{k})\Gamma(1-\hat{k})$, $\hat{\xi}=l_1+(l_2-\hat{\alpha})/\hat{k}$
Log-normal	$z=\surd(8/3)\Phi^{-1}\left(\dfrac{1+t_3}{2}\right)$, $\hat{\sigma}\approx 0.999\,281z-0.006\,118z^3+0.000\,127z^5$, $\hat{\mu}=\log\{l_2/\mathrm{erf}(\sigma/2)\}-\hat{\sigma}^2/2$, $\hat{\xi}=l_1-\exp(\hat{\mu}+\hat{\sigma}^2/2)$
Gamma	(ξ known) $t=l_2/l_1$; if $0<t<\frac{1}{2}$ then $z=\pi t^2$ and $\hat{\alpha}\approx(1-0.3080z)/(z-0.058\,12z^2+0.017\,65z^3)$; if $\frac{1}{2}\leqslant t<1$ then $z=1-t$ and $\hat{\alpha}\approx(0.7213z-0.5947z^2)/(1-2.1817z+1.2113z^2)$; $\hat{\beta}=l_1/\hat{\alpha}$

†γ is Euler's constant; Φ^{-1} is the inverse standard normal distribution function.

L-moment estimators, when compared with those of maximum likelihood estimators, usually show the method of L-moments to be reasonably efficient. For example, the efficiencies of the L-moment estimators of location and scale for the normal distribution are 100% and 97.8% respectively; for the Gumbel distribution the corresponding values are 99.6% and 75.6%. Asymptotic efficiencies of L-moment estimators tend to be lower, but still reasonably high, for distributions with more than two parameters. For example, the parameters of the GEV distribution are all estimated with at least 70% efficiency when the shape parameter k satisfies $-0.2 \leqslant k \leqslant 0.2$ (Hosking *et al.*, 1985).

In practice only a finite sample is available, and asymptotic theory is not always a reliable guide to finite sample performance: Hannan (1987) refers to the 'near irrelevance of asymptotic criteria in small samples'. Hosking *et al.* (1985) compared maximum likelihood estimation and the method of L-moments for the GEV distribution, paying particular attention to estimation of quantiles in the upper tail of the distribution. For all values of the shape parameter in the range $-0.5 < k < 0.5$, and for all samples sizes up to 100, estimates obtained by the method of L-moments have lower root-mean-square error than the maximum likelihood estimates. Results for $k = -0.2$ are given in Table 3. Similar results were reported by Hosking and Wallis (1987a) for generalized Pareto distributions with shape parameters in the range $-0.5 < k \leqslant 0$.

4.2. *Hypothesis Testing*

The use of L-moments to describe the main characteristics and to estimate the parameters of probability distributions extends naturally to testing hypotheses about distributional form. For example, for the exponential distribution, Gail and Gastwirth (1978) showed that the 'Gini index', which we recognize as the sample

TABLE 3
Root-mean-square error of quantile estimators for a GEV distribution with shape parameter $k = -0.2$†

n	*Method*	*F* 0.9	*F* 0.99	*F* 0.999
25	L	0.27	0.45	0.98
	ML	0.32	0.98	>1
50	L	0.19	0.33	0.63
	ML	0.22	0.40	0.68
100	L	0.14	0.24	0.42
	ML	0.15	0.25	0.44

†L denotes the method of L-moments; ML denotes maximum likelihood. Tabulated quantities are {root-mean-square error of $\hat{x}(F)\}/x(F)$.

L-CV l_2/l_1, has mean $\frac{1}{2}$, variance $1/12(n-1)$ and an asymptotic normal distribution. Thus the statistic

$$G_n = \{12(n-1)\}^{1/2}(l_2/l_1 - \tfrac{1}{2})$$

may be used as a test of exponentiality, with critical values obtained from the standard normal distribution. Gail and Gastwirth found this test to be a powerful test of exponentiality against a variety of alternatives and recommended it as the best of those considered in their simulation study.

In a similar spirit it can be shown that for a normal population the statistic t_3 is asymptotically normally distributed with zero mean and variance $0.1866n^{-1}$. This approximate variance is a little inaccurate when n is small: unpublished simulation studies by J. R. M. Hosking and J. R. Wallis imply that a better approximation to the variance is $v_n = 0.1866n^{-1} + 0.8n^{-2}$. Thus we propose

$$N_n = v_n^{-1/2} t_3$$

as a test statistic for normality against skew alternatives; again, critical values are obtained by reference to the standard normal distribution.

Tests of parametric hypotheses can also be obtained. For example, testing whether the shape parameter k of the GEV distribution is zero may be regarded as testing whether the distribution is Gumbel, the Gumbel distribution being the special case $k = 0$ of the GEV distribution. Hosking *et al.* (1985) derived a test of this hypothesis based on the statistic $(n/0.5633)^{1/2}\hat{k}$, where $\hat{k}$ is the L-moment estimator of the GEV shape parameter k and is a function of the statistic $\tilde{\tau}_3$ based on the plotting position $(i-0.35)/n$. Monte Carlo simulations found the performance in small samples of the L-moment test to be almost equal to that of the modified likelihood ratio test recommended by Hosking (1984); the L-moment test involves much less computation.

5. FINAL REMARKS

We have shown that L-moments provide a unified approach to statistical inference for complete samples from continuous univariate distributions. Some extensions of the approach may naturally be considered.

L-moments can be used with discrete distributions: definition (2.1) remains valid and the estimators l_r defined by equation (3.1) are still unbiased estimators of the λ_r. The inversion theorem, theorem 2.3, is valid for discrete random variables, provided that the quantile function is 'normalized' in the sense of Widder (1941), i.e. that

$$\lim_{\epsilon \to 0} \tfrac{1}{2}\{x(F+\epsilon) + x(F-\epsilon)\} = x(F) \qquad \text{for all } F \in (0, 1).$$

Expressions for L-moments of common discrete distributions tend, however, to be complicated. For a binomial distribution with parameters n and p, for example, expectations of order statistics have been derived by Ramasubban (1958) and imply that $\lambda_2 = np(1-p)\,{}_2F_1(1-n, \tfrac{1}{2}; 2; 4p(1-p))$, where ${}_2F_1$ denotes the hypergeometric function.

Censoring of data poses in principle no obstacle to the use of L-moments. Consider for example the sample L-moment l_2 calculated from a sample of $n-m$ observed values, m observations having been censored above a threshold. If the threshold value, x_0 say, is known, then l_2 has mean $\tfrac{1}{2}E(X_{2:2} - X_{1:2} \mid X_{2:2} \leqslant x_0)$; if the number of censored values was prespecified, then l_2 has mean

$$\frac{1}{2}\binom{m}{2}^{-1} E\left\{ \sum_{1 \leqslant i<j \leqslant m}\sum (X_{j:n} - X_{i:n}) \right\}.$$

Parameters of distributions can be estimated by equating the sample L-moments to their expected values, but the estimates are rarely so computationally simple as for complete samples.

No extension of L-moments to multivariate distributions is immediately apparent. The seemingly most promising approach would measure the association between random variables X and Y by a ratio of linear combinations of concomitants of order statistics, as defined by David (1973). Such measures have been derived by Barnett *et al.* (1976) and Schechtman and Yitzhaki (1987) but are asymmetric: the association between X and Y is not in general the same as that between Y and X.

ACKNOWLEDGEMENT

The author is grateful to the referees and the Editor for many helpful comments.

REFERENCES

von Andrae, C. C. G. (1872) Über die Bestimmung das wahrscheinlichen Fehlers durch die gegebenen Differenzen von *m* gleich genauen Beobachtungen einer Unbekannten. *Astronom. Nachtn*, **79**, 257–272.

Balanda, K. P. and MacGillivray, H. L. (1988) Kurtosis: a critical review. *Am. Statistn*, **42**, 111–119.

Barnett, V., Green, P. J. and Robinson, A. (1976) Concomitants and correlation estimates. *Biometrika*, **63**, 323–328.

Bickel, P. J. and Lehmann, E. L. (1975) Descriptive statistics for nonparametric models: I, Introduction; II, Location. *Ann. Statist.*, **3**, 1038–1069.

—— (1976) Descriptive statistics for nonparametric models: III, Dispersion. *Ann. Statist.*, **4**, 1139–1158.

Blom, G. (1980) Extrapolation of linear estimates to larger sample sizes. *J. Am. Statist. Ass.*, **75**, 912–917.

Bowley, A. L. (1937) *Elements of Statistics*, 6th edn. New York: Scribner.

Chan, L. K. (1967) On a characterization of distributions by expected values of extreme order statistics. *Am. Math. Mthly*, **74**, 950–951.

Chernoff, H., Gastwirth, J. L. and Johns, M. V. (1967) Asymptotic distribution of linear combinations of functions of order statistics with applications to estimation. *Ann. Math. Statist.*, **38**, 52–72.
Cunnane, C. (1978) Unbiased plotting positions—a review. *J. Hydrol.*, **37**, 205–222.
Dalén, J. (1987) Algebraic bounds on standardized sample moments. *Statist. Prob. Lett.*, **5**, 329–331.
David, H. A. (1968) Gini's mean difference rediscovered. *Biometrika*, **55**, 573–575.
—— (1973) Concomitants of order statistics. *Bull. Inst. Int. Statist.*, **45**, 295–300.
—— (1981) *Order Statistics*, 2nd edn. New York: Wiley.
Downton, F. (1966) Linear estimates with polynomial coefficients. *Biometrika*, **53**, 129–141.
Fraser, D. A. S. (1957) *Nonparametric Methods of Statistics*, ch. 4. New York: Wiley.
Gail, M. H. and Gastwirth, J. L. (1978) A scale-free goodness-of-fit test for the exponential distribution based on the Gini statistic. *J. R. Statist. Soc.* B, **40**, 350–357.
Gini, C. (1912) Variabilità e mutabilità, contributo allo studio delle distribuzioni e delle relazione statistiche. *Stud. Econ.-Giur. R. Univ. Cagl.*, **3**, 3–159.
Greenwood, J. A., Landwehr, J. M., Matalas, N. C. and Wallis, J. R. (1979) Probability weighted moments: definition and relation to parameters of several distributions expressable in inverse form. *Wat. Resour. Res.*, **15**, 1049–1054.
Greis, N. P. and Wood, E. F. (1981) Regional flood frequency estimation and network design. *Wat. Resour. Res.*, **17**, 1167–1177; corrections, **19** (1983), 589–590.
Groeneveld, R. A. and Meeden, G. (1984) Measuring skewness and kurtosis. *Statistician*, **33**, 391–399.
Hannan, E. J. (1987) The Cochrane and Orcutt papers. In *Specification Analysis in the Linear Model*, pp. 9–18. London: Routledge and Kegan Paul.
Harter, H. L. (1984) Another look at plotting positions. *Communs Statist. Theor. Meth.*, **13**, 1613–1633.
Hinkley, D. V. (1975) On power transformations to symmetry. *Biometrika*, **62**, 101–111.
Hoeffding, W. (1948) A class of statistics with asymptotically Normal distribution. *Ann. Math. Statist.*, **19**, 293–325.
Hosking, J. R. M. (1984) Testing whether the shape parameter is zero in the generalized extreme-value distribution. *Biometrika*, **71**, 367–374.
—— (1986) The theory of probability weighted moments. *Research Report RC12210*. IBM Research, Yorktown Heights.
—— (1989) Some theoretical results concerning *L*-moments. *Research Report RC14492*. IBM Research, Yorktown Heights.
Hosking, J. R. M. and Wallis, J. R. (1987a) Parameter and quantile estimation for the generalized Pareto distribution. *Technometrics*, **29**, 339–349.
—— (1987b) An "index-flood" procedure for regional rainfall frequency analysis. *Eos Trans. Am. Geophys. Un.*, **68**, 312.
Hosking, J. R. M., Wallis, J. R. and Wood, E. F. (1985) Estimation of the generalized extreme-value distribution by the method of probability-weighted moments. *Technometrics*, **27**, 251–261.
Kaigh, W. D. and Driscoll, M. F. (1987) Numerical and graphical data summary using *O*-statistics. *Am. Statistn*, **41**, 25–32.
Kirby, W. (1974) Algebraic boundedness of sample statistics. *Wat. Resour. Res.*, **10**, 220–222.
Konheim, A. G. (1971) A note on order statistics. *Am. Math. Mthly*, **78**, 524.
Lanczos, C. (1957) *Applied Analysis*, p. 286. London: Pitman.
Landwehr, J. M., Matalas, N. C. and Wallis, J. R. (1979a) Probability weighted moments compared with some traditional techniques in estimating Gumbel parameters and quantiles. *Wat. Resour. Res.*, **15**, 1055–1064.
—— (1979b) Estimation of parameters and quantiles of Wakeby distributions. *Wat. Resour. Res.*, **15**, 1361–1379; correction, 1672.
Locke, C. and Spurrier, J. D. (1976) The use of *U*-statistics for testing Normality against non-symmetric alternatives. *Biometrika*, **63**, 143–147.
Mallows, C. L. (1973) Bounds on distribution functions in terms of expectations of order-statistics. *Ann. Probab.*, **1**, 297–303.
Moore, D. S. (1968) An elementary proof of asymptotic normality of linear functions of order statistics. *Ann. Math. Statist.*, **39**, 263–265.
Oja, H. (1981) On location, scale, skewness and kurtosis of univariate distributions. *Scand. J. Statist.*, **8**, 154–168.
Ramasubban, T. A. (1958) The mean difference and the mean deviation of some discontinuous distributions. *Biometrika*, **45**, 549–556.

Randles, R. H. and Wolfe, D. A. (1979) *Introduction to the Theory of Nonparametric Statistics*, ch. 3. New York: Wiley.
Rao, C. R. (1973) *Linear Statistical Inference and its Applications*, 2nd edn, sect. 6g. New York: Wiley.
Schechtman, E. and Yitzhaki, S. (1987) A measure of association based on Gini's mean difference. *Communs Statist. Theor. Meth.*, **16**, 207-231.
Serfling, R. J. (1980) *Approximation Theorems of Mathematical Statistics*. New York: Wiley.
Sillitto, G. P. (1951) Interrelations between certain linear systematic statistics of samples from any continuous population. *Biometrika*, **38**, 377-382.
——— (1969) Derivation of approximants to the inverse distribution function of a continuous univariate population from the order statistics of a sample. *Biometrika*, **56**, 641-650.
State of California Department of Water Resources (1981) Rainfall depth-duration-frequency for California.
Stigler, S. M. (1974) Linear functions of order statistics with smooth weight functions. *Ann. Statist.*, **2**, 676-693; correction, **7** (1979), 466.
Tukey, J. W. (1960) The practical relationship between the common transformations of percentages of counts and of amounts. *Technical Report 36*. Statistical Techniques Research Group, Princeton University.
Wallis, J. R. (1980) Risk and uncertainties in the evaluation of flood events for the design of hydraulic structures. In *Piene e Siccità* (eds E. Guggino, G. Rossi and E. Todini), pp. 3-36. Catania: Fondazione Politecnica del Mediterraneo.
Wallis, J. R., Matalas, N. C. and Slack, J. R. (1974) Just a moment! *Wat. Resour. Res.*, **10**, 211-219.
Widder, D. (1941) *The Laplace Transform*. Princeton: Princeton University Press.
van Zwet, W. R. (1964) *Convex Transformations of Random Variables*, Mathematical Centre Tract 7. Amsterdam: Mathematisch Centrum.