A Contribution to Multivariate L-Moments: L-Comoment Matrices

Robert Serfling* and Peng Xiao†
University of Texas at Dallas

December 2006

Abstract

Multivariate statistical analysis relies heavily on moment assumptions of second order and higher. With increasing interest in heavy tailed distributions, however, it is desirable to describe dispersion, skewness, and kurtosis under merely first order moment assumptions. Here the univariate L-moments of Hosking (1990) are extended to "L-comoments" analogous to covariance. For certain models, the second order case yields correlational analysis coherent with classical correlation but also meaningful under just first moment assumptions. We develop properties and estimators for L-comoments, illustrate for several multivariate models, examine behavior of sample multivariate L-moments with heavy-tailed data, and discuss applications to financial risk analysis and regional frequency analysis.

AMS 2000 Subject Classification: Primary 62G35 Secondary 62H05.

Key words and phrases: L-moments; nonparametric; multivariate; coskewness; cokurtosis; financial risk analysis; regional frequency analysis.

1 Introduction

A present limitation of multivariate statistical analysis is heavy reliance on moment assumptions of second order and higher. With increasing interest, however, in model-

* *Corresponding author.* Department of Mathematical Sciences, University of Texas at Dallas, Richardson, Texas 75083-0688, USA. Email: `serfling@utdallas.edu`. Website: `www.utdallas.edu/~serfling`.

† Department of Mathematical Sciences, University of Texas at Dallas, Richardson, Texas 75083-0688, USA. Email: `pengx@utdallas.edu`.

ing with heavy tailed data, we would like to characterize descriptive features, typically dispersion, skewness, and kurtosis, under low order moment assumptions. Here we introduce a new multivariate analysis methodology that contributes toward this goal in both parametric and nonparametric settings.

Our approach extends the univariate "L-moments" of Hosking [16] to a notion of "L-comoments" which have interpretations similar to the classical central moment covariance, coskewness, and cokurtosis but also possess the features of L-moments, remaining well-defined for all orders under merely a first moment assumption. The multivariate extensions of L-moments for all orders higher than two are thus matrices – L-covariance, L-coskewness, L-cokurtosis, etc. In this connection, for example, under certain assumptions the corresponding L-correlation provides a coherent extension of the classical correlation to the case of only first moments (see Proposition 3).

Section 1.1 provides background and perspective, Section 1.2 introduces univariate L-moments and their attractive properties, and Section 2 presents a few technical results for L-moments needed in treating the multivariate case in Section 3. Section 4 provides illustrations and applications, Section 5 discusses further studies. A considerably more detailed version of this paper is available at www.utdallas.edu/~serfling.

1.1 Background and Perspective

For measuring descriptive features of a univariate distribution, central moments are popular but confined to sufficiently light-tailed distributions. An appealing alternative is the series of L-moments, expectations of strategically selected linear functions of order statistics, which are finite for all orders under merely a finite first moment assumption. A formal and comprehensive treatment of L-moments was first developed in [16]. Parametric fitting of distributions by a "method of L-moments", or exploratory nonparametric analysis via L-moments as descriptive measures, may be carried out. With interest in heavy-tailed distributions, extensive L-moment methodology has been developed for regional frequency analysis in environmental science [20].

It is timely to extend the notion of L-moments to the multivariate case. Except for the extension of the univariate mean to the multivariate vector mean, this has remained open for lack of a notion of linear functions of order statistics in higher dimensional space. Hosking [16, p. 122] writes that "No extension of L-moments to multivariate distributions is immediately apparent." On the other hand, he also mentions that the "seemingly most promising approach" would be to use the notion of concomitants of order statistics to measure association of random variables. The present paper develops this insight into a formal approach. For jointly distributed (V, W) with finite mean we define a notion of *L-comoment of order* k, $k \geq 2$ (for $k = 2$ the "Gini covariance" studied in [36], [43], [31]). For $\boldsymbol{X} = (X_1, \ldots, X_d)'$ in $\mathbb{R}^d$ with finite mean, for each $k \geq 2$ the corresponding *k-th multivariate L-moment* is then the $d \times d$ matrix of L-comoments of order k for the ordered pairs (X_i, X_j), $1 \leq i, j, \leq d$. These provide new descriptive tools having practical utility for all orders similar to the widely used classical covariance matrix.

1.2 Univariate L-Moments: Definitions and Features

Essential to our development is an understanding of univariate L-moments. With $X_{1:k} \leq X_{2:k} \leq \ldots \leq X_{k:k}$ denoting the ordered observations for a sample of size k from a univariate distribution, the kth L-moment is defined as

$$\lambda_k = k^{-1} \sum_{j=0}^{k-1} (-1)^j \binom{k-1}{j} E(X_{k-j:k}). \tag{1}$$

Clearly, the L-moments are *scale equivariant.* The first L-moment, the mean $\lambda_1 = E(X_{1:1})$, is *translation equivariant.* For $k \geq 2$ the L-moments are linear contrasts among expected order statistics and hence *translation invariant*:

$$\lambda_k(\theta + \eta X) = \eta\, \lambda_k(X), \tag{2}$$

for $\eta > 0$ and arbitrary θ. Also, $\lambda_k(-X) = (-1)^k \lambda_k(X)$.

The 2nd L-moment, $\lambda_2 = \frac{1}{2}E(X_{2:2} - X_{1:2})$, measures *spread* and in fact is one-half the classical *Gini mean difference* [12]. Besides its intrinsic role, λ_2 is used to define *scale-free* higher-order descriptive measures, $\tau_k = \lambda_k/\lambda_2$, $k \geq 3$, called *L-moment ratios* [20]. Very conveniently for practical use and interpretation, these satisfy [15]

$$-1 \leq \tau_k \leq 1, \ \ k \geq 3. \tag{3}$$

In comparison, the classical central moment analogues (further discussed in Section 3.1.1) do not satisfy any such inequality.

The 3rd L-moment, $\lambda_3 = \frac{1}{3}E(X_{3:3} - 2X_{2:3} + X_{1:3})$, is simply the difference in expectations of the two spacings from a sample of size 3 and hence measures *skewness* (unscaled). As pointed out in [16], by the result [33] that the expected range for sample size 3 is three-halves the expected range for sample size 2, τ_3 is a direct analogue of *Bowley's skewness measure* [4]. The 4th L-moment, $\lambda_4 = \frac{1}{4}E(X_{4:4} - 3X_{3:4} + 3X_{2:4} - X_{1:4})$, measures *kurtosis*, as argued very nicely in [16].

Attractive properties include: finite if first moment finite; distribution determined by L-moments; L-functional structure with mutually orthogonal weight functions; L-statistic and U-statistic structures; unbiasedness. For Normal(μ, σ^2), $\lambda_1 = \mu$, $\lambda_2 = \pi^{-1/2}\sigma$, $\lambda_3 = 0$, $\lambda_4 = (30\pi^{-1} \arctan \sqrt{2} - 9)\pi^{-1/2}\sigma$. For uniform$(a, b)$, $\lambda_1 = (a + b)/2$, $\lambda_2 = (b - a)/6$, and $\lambda_k = 0$, $k \geq 3$. (The zero skewness, kurtosis, etc., for uniform distributions is not shared by the central moments.) See [20] for other examples.

2 Univariate L-Moments: Technical Basics

We provide certain results for univariate L-moments, some new, that are used in Section 3. Let cdf F have quantile function F^{-1} and L-moment sequence $\{\lambda_k\}$.

2.1 Representations

Substitution into (1) of a standard expression for the expected value of an order statistic [6] yields a classical *L-functional representation* [39, Chap. 8], [17]

$$\lambda_k = \int_0^1 F^{-1}(u)\, P^*_{k-1}(u)\, du, \tag{4}$$

where $P^*_k(u) = \sum_{j=0}^{k} p^*_{k,j}\, u^j$, with $p^*_{k,j} = (-1)^{k-j}\binom{k}{j}\binom{k+j}{j}$. The orthogonal polynomials $P^*_r(u)$, $0 \leq u \leq 1$, $r = 0, 1, 2, \ldots$, comprise the *shifted Legendre* system. By the orthogonality, the λ_k capture differing types of information about F. For detailed discussion see [16] and [20, §§2.4–2.5].

Straightforward transformation in (4) using $P^*_0(u) \equiv 1$ and the orthogonality leads to a representation *in terms of covariance*:

$$\lambda_k = \begin{cases} E(X), & k = 1; \\ \mathrm{Cov}(X, P^*_{k-1}(F(X))), & k \geq 2. \end{cases} \tag{5}$$

For $k \geq 2$, equation (5) facilitates an illuminating characterization: the kth L-moment is the covariance of X and a particular function of its *rank* $F(X)$. In particular,

$$\lambda_2 = 2\,\mathrm{Cov}(X, F(X)) = \mathrm{Cov}(X, 2F(X) - 1) \tag{6}$$

exhibits λ_2 as the covariance of X and its *centered rank*, a well-known representation [40] for the Gini mean difference. By Cauchy-Schwarz we obtain a comparison of the second L-moment with the usual standard deviation:

$$\lambda_2 \leq \sigma/\sqrt{3}, \tag{7}$$

given in [32] and equivalently in [35]. This will be used in Section 4.3. For $k = 3$, we have $\lambda_3 = -6\,\mathrm{Cov}(X, F(X)(1 - F(X)))$, the covariance of X and a function symmetric about the median of F, and it follows that λ_3 is zero for F symmetric.

It is also readily derived that the kth L-moment has a representation as *the expected value of an L-statistic*:

$$\lambda_k = n^{-1} \sum_{r=1}^{n} w_{r:n}^{(k)}\, E(X_{r:n}), \tag{8}$$

where $w_{r:n}^{(k)} = \sum_{j=0}^{\min\{r-1,k-1\}} (-1)^{k-1-j} \binom{k-1}{j} \binom{k-1+j}{j} \binom{n-1}{j}^{-1} \binom{r-1}{j}$.

2.2 Estimation

The sample version [20, formula (2.59)] of (8) is

$$\widehat{\lambda}_k = n^{-1} \sum_{r=1}^{n} w_{r:n}^{(k)}\, X_{r:n}, \tag{9}$$

– an *L-statistic* in form and *unbiased*. For $k = 1$ and 2, (9) yields $\widehat{\lambda}_1 = \overline{X}$ and (see [39, p. 263] and [20, p. 30]) $\widehat{\lambda}_2 = \frac{1}{2} G$, where $G = \binom{n}{2}^{-1} \sum_{1 \le i < j \le n} |X_i - X_j|$, the U-statistic known as Gini's mean difference. In fact, *each* $\widehat{\lambda}_k$ is a U-statistic. To see this, note from (1) that $\lambda_k = E(h(X_1, \ldots, X_k))$, where $h(x_1, \ldots, x_k) = k^{-1} \sum_{j=0}^{k-1} (-1)^j \binom{k-1}{j} x_{k-j:k}$. Now, for a kernel $h(x_1, \ldots, x_k)$ that is a linear combination of the order statistics of its arguments, it follows by a straightforward derivation or by a technique of Blom [3] that the corresponding U-statistic based on a sample of size n may be expressed as a linear combination of the order statistics of the full sample. Consequently, the U-statistic based on $h(x_1, \ldots, x_k)$ agrees with the L-statistic (9).

Under suitable second moment conditions, standard theory for U- and L-statistics [39, Chaps. 5 and 8] yields that *the vector of the first k L-moments is asymptotically k-variate normal*. These and related results are given in [16].

3 L-Comoments and Multivariate L-Moments

We now introduce *L-comoments* and examine properties, inequalities, representations, estimators, and convergence. We conclude with L-comoment matrices.

3.1 Definition and Properties of L-Comoments

3.1.1 Preliminary on Central Comoments

Let $(X^{(1)}, X^{(2)})$ have cdf F with marginal distributions F_1 and F_2, means μ_1 and μ_2, central moments $\mu_k^{(1)}$ and $\mu_k^{(2)}$, $2 \le k \le K$, and (scaled) central moment coefficients $\psi_k^{(i)} = \mu_k^{(i)}/(\mu_2^{(i)})^{k/2}$, $3 \le k \le K$. The $\psi_k^{(i)}$, $k \ge 3$, do not satisfy any universal bounds and can have arbitrarily large magnitudes, so that interpretation of sample values is by comparison with values from specific reference distributions.

Related *comoments* are (asymmetric) higher order analogues of covariance that have been developed in financial risk modeling [34]. For $k \ge 2$, the *kth central comoment of* $X^{(1)}$ *with respect to* $X^{(2)}$ is defined as $\xi_{k\,[12]} = \mathrm{Cov}(X^{(1)}, (X^{(2)} - \mu_1^{(2)})^{k-1})$. (The asymmetric counterpart is denoted $\xi_{k\,[21]}$.) For 2nd order, $\xi_{2\,[12]} = \xi_{2\,[21]} = \sigma_{12}$, the usual *covariance*, whose symmetry is merely an artifact of the definition rather than a feature necessarily desired for comoments in general. For higher order cases one could produce symmetric versions by taking signed versions of $\sqrt{\xi_{k\,[12]}\,\xi_{k\,[21]}}$, for example, but ordered pairs $(\xi_{k\,[12]}, \xi_{k\,[21]})$, $k \ge 3$, carry greater information while still being simple and therefore are preferred. Drawing upon familiarity with covariance, it is straightforward to interpret the higher order central comoments. For example, the *coskewness* $\xi_{3\,[12]}$ of $X^{(1)}$ with respect to $X^{(2)}$ increases or decreases with relatively higher or lower weight on points $(x^{(1)}, x^{(2)})$ with positive deviations $x^{(1)} - \mu_1^{(1)}$, for given squared deviation $(x^{(2)} - \mu_1^{(2)})^2$. Scale-free versions are given by $\psi_{k\,[12]} = \xi_{k\,[12]}/(\mu_2^{(1)})^{1/2}(\mu_2^{(2)})^{(k-1)/2}$, for $k = 2$ the usual correlation coefficient ρ_{12}.

3.1.2 L-Comoments

Now take $(X^{(1)}, X^{(2)})$ having cdf F with finite mean, marginals F_1 and F_2, and L-moment sequences $\{\lambda_k^{(1)}\}$ and $\{\lambda_k^{(2)}\}$. By analogy with the covariance representation (5) for L-moments, and also by analogy with the central comoments, we define the *kth L-comoment of* $X^{(1)}$ *with respect to* $X^{(2)}$ by

$$\lambda_{k\,[12]} = \mathrm{Cov}(X^{(1)}, P_{k-1}^{*}(F_2(X^{(2)}))) \tag{10}$$

with asymmetric counterpart $\lambda_{k\,[21]}$. It is readily checked that $\lambda_{k\,[12]}$ is *translation invariant* and *scale equivariant* with respect to transformations of $X^{(1)}$ and *translation and scale invariant* with respect to transformations of $X^{(2)}$:

$$\lambda_{k\,[12]}(\theta + \eta X^{(1)}, \zeta + \beta X^{(2)}) = \eta\, \lambda_{k\,[12]}(X^{(1)}, X^{(2)}), \tag{11}$$

for positive η and β and arbitrary θ and ζ. Appropriate scaled versions are given by the *L-comoment coefficients* $\tau_{k\,[12]} = \lambda_{k\,[12]}/\lambda_2^{(1)}$, analogues of the τ_k. Here $\tau_{2\,[12]}$ is the *L-correlation* of $X^{(1)}$ with respect to $X^{(2)}$, also denoted $\rho_{[12]}$.

As with central comoments, symmetric L-comoments are possible [43], but the more fundamental notion of an ordered pair of asymmetric comoments is preferred. Fortuitously, L-comoments provide this option even in the 2nd order case. Indeed, the (asymmetric) L-correlations arise naturally for a decomposition of the 2nd L-moment of a sum into a weighted sum of 2nd L-moments of the individual terms: for univariate $Y_1, \ldots, Y_n$ and $S = Y_1 + \cdots + Y_n$, we have $\lambda_2(S) = 2\,\mathrm{Cov}(S, F_S(S)) = 2\sum_{i=1}^{n} \mathrm{Cov}(Y_i, F_S(S)) = \sum_{i=1}^{n} \lambda_{2[12]}(Y_i, S) = \sum_{i=1}^{n} \rho_{[Y_i,S]}\, \lambda_2(Y_i)$.

For $X^{(1)} = X^{(2)}$, L-comoments reduce to L-moments: $\lambda_{k\,[12]} = \lambda_{k\,[21]} = \lambda_k^{(1)} = \lambda_k^{(2)}$. On the other hand, for $X^{(1)}$ and $X^{(2)}$ *independent*, $\lambda_{k\,[12]} = 0$, all $k \geq 2$.

A convenient tool is that $\lambda_{k\,[12]}$ can be expressed as the L-comoment of $E(X^{(1)} \mid X^{(2)})$ with respect to $X^{(2)}$. The following results are used to advantage in Section 4.

Proposition 1 *Let $X^{(1)}$ have finite mean. Then, for $k \geq 2$,*

$$\lambda_{k\,[12]} = \mathrm{Cov}(E(X^{(1)} \,|\, X^{(2)}), P^*_{k-1} \circ F_2(X^{(2)})) \tag{12}$$

and, under finiteness of the kth moment of $X^{(2)}$,

$$\xi_{k\,[12]} = \mathrm{Cov}(E(X^{(1)} \,|\, X^{(2)}), (X^{(2)} - \mu_1^{(2)})^{k-1}). \tag{13}$$

Proof. Using $E(X^{(1)}Q(X^{(2)})) = E(E(X^{(1)} \,|\, X^{(2)})Q(X^{(2)}))$, etc., we obtain

$$\mathrm{Cov}(X^{(1)}, Q(X^{(2)})) = \mathrm{Cov}(E(X^{(1)} \,|\, X^{(2)}), Q(X^{(2)})).$$

Now take in turn $Q(x) = P^*_{k-1} \circ F_2(x)$ and $Q(x) = (x - \mu_1^{(2)})^{k-1}$ with $k \geq 2$. $\square$

Corollary 2 *Let $X^{(1)}$ have finite mean and linear regression on $X^{(2)}$: $E(X^{(1)} \,|\, X^{(2)}) = a + b\,X^{(2)}$. Then, for $k \geq 2$,*

$$\lambda_{k\,[12]} = b\,\lambda_k^{(2)} \tag{14}$$

and, under finiteness of the kth moment of $X^{(2)}$,

$$\xi_{k\,[12]} = b\,\mu_k^{(2)}. \tag{15}$$

When $X^{(1)}$ has linear regression on $X^{(2)}$ and F_1 and F_2 are affinely equivalent, there hold simple expressions for $\tau_{k\,[12]}$ in terms of $\tau_k^{(1)}$ and $\psi_{k\,[12]}$ in terms of $\psi_k^{(1)}$. For $k = 2$ these yield that, under the assumed conditions, the L-correlation $\rho_{[12]}$ not only agrees with the classical Pearson product-moment correlation ρ_{12} but also assumes the same formula in terms of model parameters while remaining well-defined under lesser moment assumptions.

Proposition 3 *Assume (i) $(X^{(1)}, X^{(2)})$ has joint distribution with linear regression of $X^{(1)}$ on $X^{(2)}$: for some constants a and b, $E(X^{(1)} \,|\, X^{(2)}) = a + bX^{(2)}$. Also, assume* *(ii) the respective marginals F_1 and F_2 are affinely equivalent: for some constants θ and η, $F_2(x) = F_1(\eta^{-1}(x-\theta))$, i.e., $X^{(2)} \stackrel{d}{=} \theta + \eta X^{(1)}$. Then*

$$\rho_{[12]} = b\eta = \rho_{12} \tag{16}$$

holds under second moment assumptions, with the first equality valid as well under only first moment assumptions. Also, for $k \geq 2$,

$$\lambda_{k\,[12]} = b\,\eta\,\lambda_k^{(1)} = \rho_{[12]}\,\lambda_k^{(1)} \tag{17}$$

and thus

$$\tau_{k\,[12]} = \rho_{[12]}\,\tau_k^{(1)}, \tag{18}$$

and, under finiteness of the kth moment of $X^{(2)}$,

$$\xi_{k\,[12]} = b\,\eta^k\,\mu_k^{(1)} \tag{19}$$

and thus

$$\psi_{k\,[12]} = b\,\eta\,\psi_k^{(1)} = \rho_{12}\,\psi_k^{(1)}. \tag{20}$$

Proof. Under 1st moment assumptions, $\rho_{[12]} = \lambda_{2\,[12]}/\lambda_2^{(1)} = b\lambda_2^{(2)}/\lambda_2^{(1)} = b\eta$, the 1st equality by definition, the 2nd by Corollary 2 using (i) and (14), and the 3rd by (ii) and (2). Also, under 2nd moment assumptions, $\rho_{12} = \sigma_{12}/\sigma_1\sigma_2 = b\sigma_2^2/\sigma_1\sigma_2 = b\sigma_2/\sigma_1$ $= b\eta$, by (i) and (ii). This yields (16) and similar arguments yield (17) and (19). □

3.1.3 Key Inequalities for L-Comoments

Here we rigorously establish that L-correlation like the Pearson version takes values between ±1 (for previous treatments see [36], [37]). While ρ_{12} attains ±1 only under linear relationships, $\rho_{[12]}$ does so under strictly monotone relationships. In the same sense that ρ_{12} measures linearity, we thus consider $\rho_{[12]}$ to measure monotonicity.

Proposition 4 *In general,*

$$|\lambda_{2\,[12]}| = 2\,|\mathrm{Cov}(X^{(1)}, F_2(X^{(2)})| \leq 2\,\mathrm{Cov}(X^{(1)}, F_1(X^{(1)}) = \lambda_2^{(1)} \tag{21}$$

and thus

$$-1 \leq \rho_{[12]} \leq 1. \tag{22}$$

The upper (lower) bound in (22) is attained when $X^{(1)}$ and $X^{(2)}$ are related a.s. through a strictly increasing (decreasing) function, and in the case of continuous distributions this condition is necessary as well.

Proof. Let $X^{(1)}$ and $X^{(2)}$ have joint distribution F_{12}. For V and W with finite $E|V|$, $E|W|$, and $E|VW|$, we have (see [14], [25]) $\mathrm{Cov}(V, W) = \int\int[F_{V,W}(v, w) - F_V(v)F_W(w)]\,dv\,dw$. Transforming by $v = x^{(1)}$ and $w = F_2(x^{(2)})$ and checking that $F_W(w) = F_2(x^{(2)})$ and $F_{V,W}(v, w) = F_{12}(x^{(1)}, x^{(2)})$, we obtain

$$\mathrm{Cov}(X^{(1)}, F_2(X^{(2)})) = \int\int[F_{12}(x^{(1)}, x^{(2)}) - F_1(x^{(1)})F_2(x^{(2)})]\,dx^{(1)}\,dF_2(x^{(2)}). \tag{23}$$

With F_X and F_Y specified, a joint distribution $F_{X,Y}(x, y)$ satisfies the Fréchet bounds [11], [24] $\max\{F_X(x) + F_Y(y) - 1, 0\} \leq F_{X,Y}(x, y) \leq \min\{F_X(x), F_Y(y)\}$. The upper (or lower) bound is attained if $Y = g(X)$ a.s. for strictly increasing (or decreasing) function g, since then $F_{g^{-1}(Y)}(g^{-1}(Y)) = F_Y(Y)$ (or $1 - F_Y(Y)$). For continuous F_X and F_Y this condition is necessary [38, Theorem 2]. Applying the upper Fréchet bound with (23), we obtain $\mathrm{Cov}(X^{(1)}, F_2(X^{(2)})) = \int\int[\min\{F_1(x), u\} - F_1(x)\,u]\,dx\,du$. Also, hypothetically taking $X^{(2)} = X^{(1)}$, in which case $F_2(X^{(2)}) = F_1(X^{(1)})$ and the joint distribution of $X^{(1)}$ and $F_1(X^{(1)})$ attains the upper bound, the same steps yield $\mathrm{Cov}(X^{(1)}, F_1(X^{(1)})) = \int\int[\min\{F_1(x), u\} - F_1(x)\,u]\,dx\,du$. Hence $\mathrm{Cov}(X^{(1)}, F_2(X^{(2)})) \leq \mathrm{Cov}(X^{(1)}, F_1(X^{(1)}))$. Now using $\max\{a+b-1, 0\} - ab = -[\min\{1-a, b\} - (1-a)b]$, along with the lower Fréchet bound, a similar derivation leads to $\mathrm{Cov}(X^{(1)}, F_2(X^{(2)})) \geq -\mathrm{Cov}(X^{(1)}, F_1(X^{(1)}))$, completing the proof. □

Remark. Via $\int\int[\min\{F_1(x),u\}-F_1(x)\,u]\,dx\,du = \frac{1}{2}\int F_1(x)[1-F_1(x)]\,dx$, we thus have $\lambda_2 = \int F(x)[1-F(x)]\,dx$, also given in [15] and equivalently in [26]. □

Generalization of Proposition 4 to higher order comoment coefficients is somewhat problematic. The assumptions of Proposition 3, however, yield a useful result.

Corollary 5 *Under the conditions of Proposition 3, we have, for* $k \geq 2$, $|\tau_{k\,[12]}| \leq |\tau_k^{(1)}|$ *(*≤ 1*, by (3)) and* $|\psi_{k\,[12]}| \leq |\psi_k^{(1)}|$ *(*$\leq \infty$*).*

3.1.4 L-Correlation, L-Coskewness, and L-Cokurtosis

The 2nd L-comoments and the L-correlations have been studied in [36], [37], [43], and [31] as *Gini covariances* and *Gini correlations*, with emphasis on "Gini regression analysis" and applications in economics.

One way to interpret $\lambda_{2\,[12]}$ relative to σ_{12} is through the following analogue of (6):

$$\lambda_{2\,[12]} = 2\,\mathrm{Cov}(X^{(1)}, F_2(X^{(2)})) = 2\,\mathrm{Cov}(X^{(1)}, F_2(X^{(2)}) - 1/2).$$

Thus $\lambda_{2\,[12]}$ differs from σ_{12} simply in replacing the deviation $X^{(2)} - \mu_1^{(2)}$ of $X^{(2)}$ from its mean by the deviation $F_2(X^{(2)}) - 1/2$, a scale-free measure of the deviation of $X^{(2)}$ from its *median*. Similarly, for $\lambda_{3\,[12]}$ we obtain $\lambda_{3\,[12]} = \mathrm{Cov}(X^{(1)}, P_2^*(F_2(X^{(2)})))$ $= 6\,\mathrm{Cov}(X^{(1)}, (F_2(X^{(2)}) - 1/2)^2)$. Thus $\lambda_{3\,[12]}$ differs from its central analogue $\xi_{3\,[12]}$ simply by replacing the deviation $X^{(2)} - \mu_1^{(2)}$ by the deviation $F_2(X^{(2)}) - 1/2$. For $\lambda_{4\,[12]}$, we obtain $\lambda_{4\,[12]} = \mathrm{Cov}(X^{(1)}, 20(F_2(X^{(2)}) - 1/2)^3 - 3(F_2(X^{(2)}) - 1/2))$. Again the L-comoment replaces $X^{(2)} - \mu_1^{(2)}$ by $F_2(X^{(2)}) - 1/2$, except that here in addition the particular function of the deviation also changes (slightly).

3.2 Representations for $\lambda_{k\,[12]}$ in Terms of Concomitants

Consider now a sample $\{(X_i^{(1)}, X_i^{(2)}), 1 \leq i \leq n\}$ from $F(x^{(1)}, x^{(2)})$ with marginals F_1 and F_2. Corresponding to the ordered $X^{(2)}$-values $X_{1:n}^{(2)} \leq X_{2:n}^{(2)} \leq \cdots \leq X_{n:n}^{(2)}$, we call the element of $\{X_1^{(1)}, \ldots, X_n^{(1)}\}$ that is paired with $X_{r:n}^{(2)}$ the *concomitant* $X_{[r:n]}^{(12)}$ of $X_{r:n}^{(2)}$ (see [41] and [6]). It is quickly seen that $E(X_{[r:n]}^{(12)}) = n\,E(X_1^{(1)} \mid X_1^{(2)} = X_{r:n}^{(2)})$, leading immediately to $E(X_{[r:n]}^{(12)}) = n\binom{n-1}{r-1}E(X^{(1)}\,[F_2(X^{(2)})]^{r-1}\,[1 - F_2(X^{(2)})]^{n-r})$. This may be used to establish the following representation expressing L-comoments in terms of expected values of concomitants in exactly the same way that L-moments are defined in terms of expected values of order statistics. (This does not quite mean, however, that the L-comoments can be called the L-moments of the concomitants.)

Proposition 6 *The kth L-comoment of $X^{(1)}$ with respect to $X^{(2)}$ may be represented as*

$$\lambda_{k\,[12]} = k^{-1} \sum_{j=0}^{k-1} (-1)^j \binom{k-1}{j} E(X_{[k-j:k]}^{(12)}). \tag{24}$$

Proposition 6 immediately yields another proof, communicated by Jon Hosking, of the inequality (21). We merely apply to (24) the well-known result [13, Theorem 368] that, given an ordered sequence $a_1 \le a_2 \le \cdots \le a_I$ and any other sequence $b_1, \ldots, b_I$, *the sum of products $\sum_{i=1}^{I} a_i b_{\sigma(i)}$ for a permutation $(\sigma(1), \ldots, \sigma(I))$ of $(1, \ldots, I)$ attains its maximum (minimum) possible value when the sequence $b_{\sigma(i)}, \ldots, b_{\sigma(I)}$ is increasing (decreasing).*

The main role of Proposition 6, however, is to make it straightforward to obtain key results for L-comoments as analogues of those for L-moments, with concomitants in place of order statistics. For example, we obtain for a sample of size n a direct analogue of (8) and thus the basis for unbiased estimation of comoments:

Proposition 7 *For $k \ge 2$, and with $w_{r:n}^{(k)}$ the same as in (8),*

$$\lambda_{k\,[12]} = n^{-1} \sum_{r=1}^{n} w_{r:n}^{(k)}\, E(X_{[r:n]}^{(12)}). \tag{25}$$

3.3 Estimation of L-Comoments

Proposition 7 yields for the kth L-comoment the unbiased estimator

$$\widehat{\lambda}_{k\,[12]} = n^{-1} \sum_{r=1}^{n} w_{r:n}^{(k)} X_{[r:n]}^{(12)}, \tag{26}$$

an *L-statistic in the concomitants.* Further, each $\widehat{\lambda}_{k\,[12]}$ is a *U-statistic*, via $\lambda_{k\,[12]} = E(h^{(k)}((X_1^{(1)}, X_1^{(2)}), \ldots, (X_k^{(1)}, X_k^{(2)})))$, with

$$h^{(k)}((x_1^{(1)}, x_1^{(2)}), \ldots, (x_k^{(1)}, x_k^{(2)})) = k^{-1} \sum_{j=0}^{k-1} (-1)^j \binom{k-1}{j} x_{[k-j:k]}^{(12)}.$$

For $k = 2$ these L- and U-statistic representations are $\widehat{\lambda}_{2\,[12]} = n^{-1} \sum_{r=1}^{n} \frac{2r-n-1}{n-1} X_{[r:n]}^{(12)} = \binom{n}{2}^{-1} \sum_{1 \le i < j \le n} (X_{[j:n]}^{(12)} - X_{[i:n]}^{(12)})/2$, analogous to expressions for the 2nd L-moment, as expected. (The present U-statistic representation, however, cannot be reexpressed as one-half the Gini mean difference of the concomitants, because the relevant kernel in the concomitants, $(x_{[2:2]}^{(12)} - x_{[1:2]}^{(12)})/2$, is not the same as the kernel $|x_{[2:2]}^{(12)} - x_{[1:2]}^{(12)}|/2$ for the Gini mean difference.)

The asymptotic distribution of a vector of L-comoment estimators follows from standard theory for U-statistics [39]. Defining

$$g^{(i)}(x^{(1)}, x^{(2)}) = i\, E(h^{(i)}((x^{(1)}, x^{(2)}), (X_2^{(1)}, X_2^{(2)}), \ldots, (X_i^{(1)}, X_i^{(2)})))$$

and $\zeta_{ij} = \mathrm{Cov}(g^{(i)}(X^{(1)}, X^{(2)}), g^{(j)}(X^{(1)}, X^{(2)}))$, $2 \le i, j \le k$, we have

Proposition 8 *Under second moment assumptions on $X^{(1)}$, for $k \ge 2$ the vector of sample L-comoments $(\widehat{\lambda}_{2\,[12]}, \ldots, \widehat{\lambda}_{k\,[12]})'$ is asymptotically $(k-1)$-variate normal with mean $(\lambda_{2\,[12]}, \ldots, \lambda_{k\,[12]})'$ and covariance matrix $[\zeta_{ij}]/n$.*

Alternatively, this follows using (26) with results of Yang [42]. Asymptotic normality of the vector of scaled versions $\widehat{\tau}_{i\,[12]}$, $2 \le i \le k$, follows by standard results on transformations of asymptotically normal vectors.

3.4 Multivariate L-Moments

We now define "multivariate L-moments". For random d-vector $\boldsymbol{X} = (X^{(1)}, \ldots, X^{(d)})'$, the first order multivariate L-moment is simply the *vector mean* $\boldsymbol{\lambda}_1 = E(\boldsymbol{X})$, assumed finite. For $k \geq 2$, the *kth multivariate L-moment* is the matrix of kth L-comoments for all pairs $(X^{(i)}, X^{(j)})$, $1 \leq i, j \leq d$: $\boldsymbol{\Lambda}_k = (\lambda_{k\,[ij]})_{d \times d}$, with $\boldsymbol{\Lambda}_2$, $\boldsymbol{\Lambda}_3$, and $\boldsymbol{\Lambda}_4$ the *L-covariance*, *L-coskewness*, and *L-cokurtosis* matrices, respectively. Corresponding versions with scaled elements are given by $\boldsymbol{\Lambda}_k^* = (\tau_{k\,[ij]})$, the *L-comoment coefficient matrices*. The diagonals of $\boldsymbol{\Lambda}_k$ and $\boldsymbol{\Lambda}_k^*$ are the componentwise univariate L-moments and L-moment coefficients, respectively. In the illustrations of Section 4, we compare with the corresponding *central* versions, denoted $\boldsymbol{\Xi}_k = (\xi_{k[ij]})$ and $\boldsymbol{\Xi}_k^* = (\psi_{k[ij]})$, respectively, $k \geq 2$ ($\boldsymbol{\Xi}_2$ and $\boldsymbol{\Xi}_2^*$ being the usual covariance and correlation matrices).

As with classical correlation, by Proposition 4 the pairwise L-correlations are assessed through comparison with the values ± 1. No such guideline exists in the case of higher orders, however, neither for central comoment nor L-comoment coefficients, nor for the univariate central counterparts. One compensating approach is to rely upon suitable reference multivariate distributions as benchmarks. Under certion assumptions, however, which may be verified for a particular model or assumed in a nonparametric formulation, we can indeed state upper and lower bounds for L-comoment coefficients of all orders.

Proposition 9 *Assume that the components of $\boldsymbol{X} = (X^{(1)}, \ldots, X^{(d)})'$ have affinely equivalent marginal distributions and pairwise linear regressions, in the sense of the conditions of Proposition 3. Then marginal L-moment coefficients agree and likewise* *for marginal central moment coefficients:*

$$\tau_k^{(1)} = \ldots = \tau_k^{(d)} = \tau_k, \text{ say,} \tag{27}$$

$$\psi_k^{(1)} = \ldots = \psi_k^{(d)} = \psi_k, \text{ say,} \tag{28}$$

for $k \geq 3$. *Further,*

$$\rho_{[ij]} = \rho_{ij} = \rho_{[ji]}, \; 1 \leq i, j \leq d,$$

yielding, with $\mathbf{C} = (\rho_{ij}) = (\rho_{[ij]})$,

$$\Lambda_k^* = \tau_k \, \mathbf{C}, \tag{29}$$

$$\Xi_k^* = \psi_k \, \mathbf{C}. \tag{30}$$

This result follows readily from Proposition 3. In each of (29) and (30), the comoment coefficient matrix is simply the product of the univariate moment coefficient of the same order and the correlation matrix $\mathbf{C}$. The central comoment and L-comoment coefficient matrices are both, in this instance, equivalent in structure to the usual correlation matrix, which thus contains all of the multivariate shape information (in the scale-free sense).

We note an interesting open characterization issue. The univariate L-moments determine F in the case of finite mean [5]. In the multivariate case we ask, for example, to what extent the L-moments and L-comoments together determine the bivariate distributions.

4 Illustrations and Applications

In the multivariate case, tractable distributions are fewer and parametric approaches more limited than in the univariate setting. Although univariate L-moments provide a useful alternative to the classical method of moments in parametric model-fitting, and this indeed extends to the multivariate case, the widest and most significant role of multivariate L-moments lies in providing attractive *nonparametric multivariate estimators and descriptive measures.* Using the estimators and theory of Sections 2.2 and 3.3, one may readily compute for a data set sample versions of $\mathbf{\Lambda}_2^*$, $\mathbf{\Lambda}_3^*$, and $\mathbf{\Lambda}_4^*$ and (under second moment assumptions) characterize asymptotic distributions.

In Sections 4.1, 4.2, and 4.3 we illustrate for the normal, Pareto, and Farlie-Gumbel-Morgenstern multivariate distributions, the first two governed by Proposition 9, the third not. Sections 4.4 and 4.5 indicate the role of multivariate L-moments in portfolio risk analysis and regional frequency analysis.

4.1 The Multivariate Normal Distribution

For a d-variate normal model with variances σ_i^2 and covariances σ_{ij}, the assumptions of Propositions 1, 3, and 9 are fulfilled with $b = \sigma_{ij}/\sigma_j^2$, $\eta = \sigma_j/\sigma_i$, and thus $b\eta = \rho_{ij}$. The comoments are given by $\lambda_{k\,[ij]} = (\sigma_{ij}/\sigma_j^2)\lambda_k^{(j)}$ and $\xi_{k\,[ij]} = (\sigma_{ij}/\sigma_j^2)\mu_k^{(j)}$, and the comoment coefficients by $\tau_{k\,[ij]} = \rho_{ij}\tau_k$ and $\psi_{k\,[ij]} = \rho_{ij}\psi_k$, $k \geq 2$. For odd $k \geq 3$, these quantities are all 0. For even k, the central moment coefficients are invariant over parameters and readily found to be $\psi_k = (k-1)(k-3)\cdots 3\cdot 1$. The quantities τ_k are more elusive, explicit expressions for the expected values of order statistics for normal samples in terms of elementary functions being known only for sample sizes ≤ 5. For a range of larger sample sizes, however, these expected values have been computed numerically and tabulated, and approximations are available for indefinitely large sample sizes [22, pp. 94–96]. In particular, the 2nd, 3rd, and 4th normal L-moments mentioned in Section 1.2 yield $\tau_3 = 0$, and $\tau_4 = (30\pi^{-1}\arctan\sqrt{2} - 9)$.

4.2 A Multivariate Pareto Distribution

We consider here the Type II version of the multivariate Pareto distribution of [2], given by the d-variate joint cdf

$$F(x^{(1)}, \ldots, x^{(d)}) = 1 - \left[1 + \sum_{i=1}^{d} \left(\frac{x^{(i)} - \theta_i}{\sigma_i}\right)\right]^{-\alpha}, \qquad (31)$$

for $x^{(i)} > \theta_i$ and $\sigma_i > 0$, $1 \le i \le d$, and $\alpha > 0$. The kth moment is finite if $k < \alpha$. Many typical applications involve heavy-tailed modeling, with α in the range 1 to 2 for quite diverse data sets (see, for example, [2, Appendix A], [22, p. 575], and [27]). With $\theta_i = \sigma_i$, $1 \le i \le d$, (31) has long been used in actuarial science and economics. With $\theta_i \equiv 0$, (31) arises in reliability theory Nayak [30]. For general discussion of model (31), see [2] and [24, pp. 380–382 and 603–605].

For *parametric* inference using this model, one may use the maximum likelihood method, the classical method-of-moments, or the analogous method-of-L-moments. We describe below the derivation of tractable formulas for all the relevant L-moments, L-comoments, central moments, central comoments, and related coefficients.

We also apply model (31) to explore, comparatively with central versions, the empirical behavior of the sample L-moments, L-comoments, and related coefficients as *nonparametric descriptive measures* based on data from an unknown and possibly heavy-tailed distribution. Some sampling and simulation results are provided below.

4.2.1 Formulas

For $\boldsymbol{X} = (X^{(1)}, \ldots, X^{(d)})'$ having distribution (31), $X^{(i)}$ has marginal distribution $F_i(x^{(i)}) = 1 - [1 + (\sigma_i^{-1}(x^{(i)} - \theta_i))]^{-\alpha}$ and linear regression on $X^{(j)}$, fulfilling the assumptions and conclusions of Propositions 1, 3, and 9 with $\eta = \sigma_j/\sigma_i$, $b = \sigma_i/\sigma_j\alpha$, and $\boldsymbol{C} = (c_{ij})$ with $c_{ij} = 1$ or α^{-1} for $i = j$ or $i \neq j$. For F_i we obtain $\lambda_1^{(i)} =$

$\theta_i + \sigma_i/(\alpha - 1) = \mu_1^{(i)}$, $\lambda_k^{(i)} = \sigma_i\, \alpha \prod_{j=0}^{k-2}(j\alpha + 1)/\prod_{j=1}^{k}(j\alpha - 1)$, for $k \geq 2$, and thus $\tau_k = \prod_{j=0}^{k-2}(j\alpha + 1)/\prod_{j=3}^{k}(j\alpha - 1)$, for $k \geq 3$. For computation of the kth central moment, we assume without loss of generality that $\theta_i = 0$ and use [2, (3.3.8)] $E(X^{(i)}) = \sigma_i^k k!/(\alpha - 1) \cdots (\alpha - k)$, yielding $\mu_k^{(i)}$ and in turn ψ_k, $2 \leq k < \alpha$. In particular, $\psi_3 = 2(\alpha + 1)((\alpha - 2)/\alpha)^{1/2}/(\alpha - 3)$ and $\psi_4 = 3(3\alpha^2 + \alpha + 2)(\alpha - 2)/\alpha(\alpha - 3)(\alpha - 4)$.

With $\mathbf{C}$, ψ_k, and τ_k as above, the comoment coefficient matrices for this model are now given by (29) and (30). Here the factors τ_k and ψ_k depend not only upon k but also upon the shape parameter α. The use of (29) requires $\alpha > 1$, while (30) requires $\alpha > k$. We thus obtain for this model an extended correlation analysis, since the formula α^{-1} for all the Pearson correlations under $\alpha > 2$ holds also for all the L-correlations under $\alpha > 1$. The maximal value 1/2 for the correlation under $\alpha > 2$ increases to 1 and becomes approached, as $\alpha \downarrow 1$.

4.2.2 Some Empirical Results

To examine the performance of sample L-moments and L-comoments, with special reference to the case of heavy-tailed data, and to compare with corresponding central versions, we provide a small simulation study using the above Pareto II model. For each of $\alpha = 1.5$, 2.5, 3.5, and 4.5, and sample sizes $n = 50$ and 500, we generated 20,000 samples from the cdf (31) with $d = 3$, $\theta_i \equiv 0$, and $\sigma_i \equiv 1$. Each trivariate observation $\boldsymbol{X} = (X^{(1)}, X^{(2)}, X^{(3)})'$ was obtained via the representation [2, p. 252] $X^{(i)} = W_i/Z$, $1 \leq i \leq 3$, with independent standard exponential random variables W_1, W_2, and W_3 and gamma$(\alpha, 1)$ random variable Z. For each sample, the L-moments, L-comoments, central moments, central comoments, and corresponding coefficients were computed for orders $k \leq 4$. With these data, we compare, on the basis of 20,000 observations each, the L-versions and central versions of *multivariate nonparametric descriptive measures for spread, skewness, and kurtosis* (taking into account that each quantity is measured in a different way by the two versions).

Selected representative results for L-moments and L-comoments of orders 2–4 as well as for L-correlation are provided for $\alpha = 1.5$ and 4.5 in Tables 1 and 2, respectively. Table 2 also includes results for central versions (which are defined for $\alpha = 4.5$). For each target parameter and sample size, we list the population value and, based on the 20,000 sample estimates, the mean (mean), median (med), coefficient of variation (CV), and relative interquartile range (RIQR, defined as IQR/med), of the estimates. The results in the tables support a number of conclusions: 1. The CV and RIQR variability measures decrease as sample size n increases. However, for $\alpha = 1.5$, the decrease in CV is only slight, reflecting higher sensitivity to extreme observations. 2. The CV and RIQR measures both increase as the order k increases, with the increase in CV for the central versions very dramatic. 3. For $\alpha = 4.5$, the L-versions are much more stable and efficient than the central versions as estimators of their respective parameters. For order ≥ 3, the central versions are especially erratic. 4. For estimation of correlation, $\alpha = 4.5$ the sample L-correlation and Pearson correlation are both fairly strong, with the L-version distinctly more stable and efficient. In the very heavy-tailed case of $\alpha = 1.5$, however, the sample L-correlation is noticeably less efficient and the sample Pearson correlation meaningless (figures not included). 5. The sample L-comoments for $\lambda_{2[12]}$ and $\lambda_{2[21]}$ (which are equal in the present model) behave very consistently for each case of α. 6. *Summary Comment.* For nonparametric moment-based description with data from a possibly heavy-tailed distribution, L-versions offer clear advantages over central versions. The gain increases with increasing order of moments and with increasing heaviness of tails.

4.3 Multivariate Farlie-Gumbel-Morgenstern Distributions

An appealing structure for joint distributions having given marginals was introduced in [28] and [10], with considerable further development leading to so-called Farlie-Gumbel-Morgenstern (FGM) classes of distributions. Here we consider [24, (44.73)]

$$F(x^{(1)}, \ldots, x^{(d)}) = \prod_{i=1}^{d} F_i(x^{(i)}) \times \left[1 + \left(\sum_{1 \le i_1 < i_2 \le d} \alpha_{i_1 i_2} (1 - F_{i_1}(x^{(i_1)}))(1 - F_{i_2}(x^{(i_2)})) \right) + \cdots + \left(\alpha_{1\,2 \cdots d} \prod_{i=1}^{d} (1 - F_i(x^{(i)})) \right) \right], \tag{32}$$

with $a_{i_1 \cdots i_\ell}$ satisfying $1 + \sum_{1 \le i_1 < i_2 \le d} \alpha_{i_1 i_2} \varepsilon_{i_1} \varepsilon_{i_2} + \cdots + \alpha_{1\,2 \cdots d} \varepsilon_1 \cdots \varepsilon_d \ge 0$, for all cases of $\varepsilon_i = \pm 1$, a sufficient condition for $F(x^{(1)}, \ldots, x^{(d)})$ to be a nondecreasing function of its arguments. In the case of mutually independent components, $\alpha_{i_1 \cdots i_\ell} \equiv 0$. For $\boldsymbol{X} = (X^{(1)}, \ldots, X^{(d)})'$ having cdf (32), $X^{(i)}$ has marginal distribution $F_i(\cdot)$, yielding marginal L-moments and central moments. For derivation of comoments and comoment coefficients, we use the bivariate distributions $F_{ij}(x^{(i)}, x^{(j)}) = F_i(x^{(i)}) F_j(x^{(j)}) [1 + \alpha_{ij} (1 - F_i(x^{(i)}))(1 - F_j(x^{(j)}))]$, with $|\alpha_{ij}| \le 1$, from which it follows [24, p. 56] that $X^{(i)}$ has linear regression on $F_j(X^{(j)})$ with slope $b = 4\,\alpha_{ij}\,\mathrm{Cov}(X^{(i)}, F_i(X^{(i)})) = 2\,\alpha_{ij}\,\lambda_2^{(i)}$. Corollary 2 then yields $\lambda_{k\,[ij]}$ and $\xi_{k\,[ij]}$, $k \ge 2$.

Now take all F_j to be *continuous*. Then the covariance factor in $\lambda_{k\,[ij]}$ is by (5) just the kth L-moment of the uniform$(0,1)$ distribution, which equals $1/6$ for $k = 2$ and 0 for $k \ge 3$, by orthogonality of the P_ℓ^*. The central comoments $\xi_{k\,[12]}$, however, are nonzero for $k \ge 3$. For $k = 2$, $\lambda_{2\,[ij]} = \alpha_{ij}\,\lambda_2^{(i)}/3$ and $\sigma_{ij} = \xi_{2\,[ij]} = \alpha_{ij}\,\lambda_2^{(i)}\,\lambda_2^{(j)}$, with corresponding correlations $\rho_{[ij]} = \alpha_{ij}/3$ under first moment assumptions and $\rho_{ij} = \alpha_{ij} \lambda_2^{(i)} \lambda_2^{(j)} / \sigma_i \sigma_j$ under second moment assumptions. By (7) we thus obtain $|\rho_{ij}| \le |\alpha_{ij}|/3 = |\rho_{[ij]}|$. Since $|\alpha_{ij}| \le 1$, both correlations are thus no greater than

1/3 in magnitude. This weak dependence is also manifest, in a new way, by the higher-order L-comoments all being 0, similar to the case of independent variables.

4.4 Modeling for Portfolio Risk Analysis in Finance

Among approaches to portfolio optimization in finance, a central role has long been played by the Capital Asset Pricing Model (CAPM), initially involving just first and second moments but recently higher moments also. Skewness measures concern evaluation of the downside risk and asymmetric volatility of a portfolio, while spread and kurtosis measures concern volatility and uncertainty in returns. For detailed discussion, see [1], [7], and [23]. Also increasing is interest in heavy-tailed distributions in modeling stock returns, raising serious concern regarding higher moment assumptions and issues of stability and robustness associated with higher order central moments and comoments. In fact, for the *marginal* distributions of jointly distributed heavy-tailed variables in risk analysis, univariate L-moments have already been applied [19]. Such treatments now can be extended using L-comoments.

4.5 Modeling for Regional Frequency Analysis in Environmental Science

Many environmental applications involve, for each variable of interest, for example streamflow, separate series of observations taken at different measurement sites within a network. This yields for a given variable multiples samples of similar data, with possible dependence within as well as between samples. One key goal is to estimate the upper quantile corresponding to occurrence of a specified “extreme” event. In many applications the site sample sizes are too small for efficient estimation of upper quantiles, and data within a suitable region are combined through “regional frequency analysis” under effective simplifying assumptions. L-moment methods have proved effective in providing stable and reliable estimates less sensitive to model assumptions and extreme observations [20]. In particular, the network is partitioned into *approximately homogeneous regions* of sites with very similar distributions for the variable of interest. For each site, the vector of the first four sample L-moments or coefficients is obtained and "unusual" sites are identified via a suitable discordancy measure. In many situations, however, several variables of interest are measured at each site, for example, streamflow, temperature, precipitation, windspeed, etc. Instead of generating different partitions separately for each variable, with the multivariate L-moments approach one can develop an extended regional frequency analysis leading to a single partition based on all the variables considered jointly.

5 Increased Robustness and Lower Moment Assumptions

Trimmed L-Moments

A modification of L-moments to obtain more robustness and reduce moment assumptions is introduced in [8]. *Trimmed L-moments* are given by increasing the conceptual sample size for the kth L-moment from k to $k + t_1 + t_2$ and using the k order statistics remaining after trimming the t_1 smallest and t_2 largest observations in the conceptual sample. Thus λ_k given by (1) becomes replaced by $\lambda_k^{(t_1,t_2)} = k^{-1} \sum_{j=0}^{k-1} (-1)^j \binom{k-1}{j} E(X_{k+t_1-j:k+t_1+t_2})$, $k \geq 1$. Except for $(t_1, t_2) = (0, 0)$, which gives the usual L-moments, the TL-moments exist under weaker moment assumptions and eliminate the influence of the most extreme observations. The sample TL-moments do not, however, improve upon the asymptotic finite sample breakdown point, 0, of the sample L-moments. See [8], [9] and [18] for detailed development.

Our definitions of L-comoments and L-comoment coefficients carry over easily to provide analogous TL-comoments and TL-comoment coefficients. While asymptotic results are not provided in [8], for (t_1, t_2) fixed as $n \to \infty$, the asymptotic results we have stated for sample L-moments and L-comoments have similar formulations and derivations for these trimmed versions.

L-Moments on Trimmed Samples

The alternative approach of defining trimmed L-moments simply as ordinary L-moments defined on a trimmed sample is mentioned in [8] without development. This yields different versions of trimmed estimators – for example, for first order the usual trimmed mean, weighting each observation equally after trimming. For $(t_1, t_2) = (\beta n, \beta n)$ with $\beta > 0$, the breakdown point improves from 0 to β. Asymptotic normality of these sample versions follows, using the U-statistic representations noted in the present paper, from results of [21] for U-statistics defined on trimmed samples.

Quantiles instead of Expectations

An analogue of L-moments that eliminates moment restrictions entirely consists of replacing each expectation in (1) by a suitable linear combination of quantiles: $\lambda_k^{(Q)} = k^{-1} \sum_{j=0}^{k-1} (-1)^j \binom{k-1}{j} \theta_{p,\alpha}(X_{k-j:k})$, where $0 \leq \alpha \leq 1/2$, $0 \leq p \leq 1/2$, and $\theta_{p,\alpha}(X_{k-j:k}) = pF^{-1}_{X_{k-j:k}}(\alpha) + (1-2p)F^{-1}_{X_{k-j:k}}(1/2) = pF^{-1}_{X_{k-j:k}}(1-\alpha)$. See [29] for general treatment. These may be extended to define LQ-comoments and related quantities. Starting with the representation (24) of L-comoments in terms of concomitants, as given in Proposition 6, we replace expectations by quantiles to define the *kth LQ-comoment of $X^{(1)}$ with respect to $X^{(2)}$* by $\lambda^{(Q)}_{k\,[12]} = k^{-1} \sum_{j=0}^{k-1} (-1)^j \binom{k-1}{j} \theta_{p,\alpha}(X^{(12)}_{[k-j:k]})$.

Variances and Covariances of Sample Versions

Exact formulae for the variances and covariances of sample L-moments and TL-moments are developed in [8], [9]. These have the form of a weighted sum of expected values of order statistics from a conceptual sample. These can also be derived for the U-statistic representations we have noted above, via standard expressions for the variances and covariances of U-statistics. In the same fashion, exact expressions for the variances and covariances of sample TL-comoments may be obtained.

Distribution-free unbiased estimators of these variances and covariances are also provided in [8], [9]. We note that, using the weighted sum of expected values from a conceptual sample to define a corresponding kernel for a U-statistic, distribution-free unbiased estimators also are given immediately by the corresponding U-statistics.

Acknowledgments

The authors very appreciatively thank William Asquith, Nick Cox, Jon Hosking, G. L. Thompson, and three anonymous referees for extremely helpful queries and suggestions that have led to many improvements in the exposition. Support by National Science Foundation Grants DMS-0103698 and CCF-0430366 is also gratefully acknowledged.

References

[1] Adcock, C., Jurczenko, E., and Maillet, B., eds. (2005). *Multi-Moment Capital Asset Pricing Models and Related Topics*. Springer-Verlag, Berlin.

[2] Arnold, B. C. (1983). *Pareto Distributions*. International Cooperative Publishing House, Fairland, Maryland.

[3] Blom, G. (1980). Extrapolation of linear estimates to larger sample sizes. *Journal of the American Statistical Association* **75** 912–917.

[4] Bowley, A. L. (1902). *Elements of Statistics*, 2nd Edition. P. S. King, London.

[5] Chan, L. K. (1967). On a characterization of distributions by expected values of extreme order statistics. *American Mathematical Monthly* **74** 950–951.

[6] David, H. A. and Nagaraja, H. N. (2003). *Order Statistics*, 3rd Edition. Wiley, New York.

[7] Dittmar, R. (2002). Nonlinear pricing kernels, kurtosis preference, and evidence from the cross-section of equity returns. *Journal of Finance* **57** 369–403.

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

[9] Elamir, E. A. and Seheult, A. H. (2004). Exact variance structure of sample L-moments. *Journal of Statistical Planning and Inference* **124** 337–359.

[10] Farlie, D. J. G. (1960). The performance of some correlation coefficients for a general bivariate distribution. *Biometrika* **47** 307–323.

[11] Fréchet, M. (1951). Sur les tableaux de corrélation dont les marges sont données. *Annales de l'Université de Lyon, Section A, Series 3* **14** 53–57.

[12] Gini, C. (1912). Variabilità e Mutabilità. Contributio allo studio delle distribuzioni e delle relazione statistiche. *Studi Economico-Giuridici della Reale Università di Cagliari* **3** 3–159.

[13] Hardy, G. H., Littlewood, J. E. and Pólya, G. (1952). *Inequalities*, 2nd edition. University Press, Cambridge.

[14] Hoeffding, W. (1940). Masstabinvariante Korrelations-theorie. *Schriften Math. Inst. Univ. Berlin* **5** 181–233.

[15] Hosking, J. R. M. (1989). Some Theoretical Results Concerning L-Moments. *Research Report RC14492*. IBM Research Division, Yorktown Heights.

[16] Hosking, J. R. M. (1990). L-moments: Analysis and Estimation of Distributions using Linear Combinations of Order Statistics. *Journal of the Royal Statistical Society, Series B* **52** 105–124.

[17] Hosking, J. R. M. (1998). L-Estimation. In *Handbook of Statistics, Vol. 17*, N. Balakrishnan and C. R. Rao (eds.), pp. 215–235. Elsevier.

[18] Hosking, J. R. M. (2005). Some Theory and Practical Uses of Trimmed L-Moments. *Research Report RC23783*. IBM Research Division, Yorktown Heights.

[19] Hosking, J. R. M., Bonti, G., and Siegel, D. (2000). Beyond the lognormal. *Risk* **13** 59–62.

[20] Hosking, J. R. M. and Wallis, J. R. (1997). *Regional Frequency Analysis: An Approach Based on L-Moments*. Cambridge University Press.

[21] Janssen, P., Serfling, R. and Veraverbeke, N. (1987). Asymptotic normality of U-statistics based on trimmed samples. *Journal of Statistical Planning and Inference* **16** 63–74.

[22] Johnson, N. L., Kotz, S., and Balakrishnan, N. (1994). *Continuous Univariate Distributions*, Volume 1, 2nd edition. Wiley, New York.

[23] Jurczenko, E. and Maillet, B. (2005). The four-moment capital asset pricing model: some basic results. In *Multi-Moment Capital Asset Pricing Models and Related Topics*, C. Adcock, E. Jurczenko, and B. Maillet (eds.), forthcoming (72 pages). Springer-Verlag, Berlin.

[24] Kotz, S., Balakrishnan, N. and Johnson, N. L. (2000). *Continuous Multivariate Distributions, Volume 1: Models and Applications*, 2nd edition. Wiley, New York.

[25] Lehmann, E. L. (1966). Some concepts of dependence. *Annals of Mathematical Statistics* **37** 1137–1153.

[26] Lerman, R. I. and Yitzhaki, S. (1984). A note on the calculation and interpretation of the Gini index. *Economics Letters* **15** 163–168.

[27] Meerschaert, M. M. and Scheffler, H.-P. (2003). Nonparametric methods for heavy tailed vector data: A survey with applications from finance and hydrology. In *Recent advances and trends in nonparametric statistics*, M. G. Akritas and D. N. Politis (eds.), pp. 265-279. Elsevier.

[28] Morgenstern, D. (1956). Einfache Beispiele zweidimensionaler Verteilungen. *Miteilingsblatt für Mathematische Statistik* **8** 234–235.

[29] Mudholkar, G. S. and Hutson, A. D. (1998). LQ-moments: Analogues of L-moments. *Journal of Statistical Planning and Inference* **71** 191–208.

[30] Nayak, T. K. (1987). Multivariate Lomax distribution: properties and usefulness in reliability theory. *Journal of Applied Probability* **24** 170–177.

[31] Olkin, I. and Yitzhaki, S. (1992). Gini regression analysis. *International Statistical Review* **60** 185–196.

[32] Plackett, R. L. (1947). Limits of the ratio of mean range to standard deviation. *Biometrika* **34** 120–122.

[33] Robbins, H. (1944). On the expected values of two statistics. *Annals of Mathematical Statistics* **15** 321–323.

[34] Rubinstein, M. E. (1973). The fundamental theorem of parameter-preference security valuation. *Journal of Financial and Quantitative Analysis* **8** 61–69.

[35] Schucany, W. R., Parr, W. C., and Boyer, J. E. (1978). Correlation structure in Farlie-Gumbel-Morgenstern distributions. *Biometrika* **65** 650–653.

[36] Schechtman, E. and Yitzhaki, S. (1987). A measure of association based on Gini's mean difference. *Communications in Statistics – Theory and Methods* **16** 207–231.

[37] Schechtman, E. and Yitzhaki, S. (1999). On the proper bounds of the Gini correlation. *Economics Letters* **63** 133–138.

[38] Schweizer, B. and Wolff, E. F. (1981). On nonparametric measures of dependence for random variables. *Annals of Statistics* **9** 879–885.

[39] Serfling, R. (1980). *Approximation Theorems of Mathematical Statistics.* John Wiley & Sons, New York.

[40] Stuart, A. (1954). The correlation between variate values and ranks in samples from a continuous distribution. *British J. Statist. Psych.* **7** 37–44.

[41] Yang, S. S. (1977). General distribution theory of the concomitants of order statistics. *Annals of Statistics* **5** 996–1002.

[42] Yang, S. S. (1981). Linear functions of concomitants of order statistics with application to nonparametric estimation of a regression function. *Journal of the American Statistical Association* **76** 658–662.

[43] Yitzhaki, S. and Olkin, I. (1991). Concentration indices and concentration curves. In *Stochastic Orders and Decision under Risk* (K. Mosler and M. Scarsini, eds.),

pp. 380–392. IMS Lecture Notes – Monograph Series, Volume 19, Hayward, California.

Table 1: L-Moment and L-Comoment Sampling Results, $\alpha = 1.5$.

Target Parameter	True Value	Sample Values $n = 50$ mean	med	CV	RIQR	$n = 500$ mean	med	CV	RIQR
λ_2	1.50	1.47	1.09	1.78	0.67	1.52	1.32	1.70	0.32
$\lambda_{2[12]}$	1.00	0.97	0.56	2.69	1.09	1.01	0.82	2.54	0.49
$\lambda_{2[21]}$	1.00	0.96	0.60	2.93	1.08	1.00	0.82	2.04	0.49
$\rho_{[12]}$	0.67	0.56	0.58	0.37	0.47	0.63	0.63	0.15	0.19
λ_3	1.07	1.05	0.66	2.49	0.95	1.09	0.89	2.37	0.44
$\lambda_{3[12]}$	0.71	0.69	0.32	3.75	1.65	0.73	0.53	3.52	0.69
λ_4	0.86	0.83	0.45	3.11	1.21	0.87	0.67	2.95	0.56
$\lambda_{4[12]}$	0.57	0.54	0.20	4.70	2.27	0.59	0.39	4.37	0.88

Table 2: L-Moment, L-Comoment, Moment, and Comoment Sampling Results, $\alpha = 4.5$.

Target Parameter	True Value	Sample Values							
		$n = 50$				$n = 500$			
		mean	med	CV	RIQR	mean	med	CV	RIQR
λ_2	0.161	0.161	0.155	0.25	0.31	0.161	0.160	0.08	0.10
$\lambda_{2[12]}$	0.036	0.036	0.032	1.04	1.36	0.036	0.035	0.32	0.43
$\lambda_{2[21]}$	0.036	0.036	0.032	1.04	1.37	0.036	0.035	0.32	0.43
$\rho_{[12]}$	0.222	0.209	0.216	0.90	1.21	0.221	0.221	0.29	0.39
λ_3	0.071	0.071	0.065	0.45	0.54	0.071	0.070	0.14	0.18
$\lambda_{3[12]}$	0.016	0.016	0.013	1.93	2.81	0.016	0.015	0.59	0.80
λ_4	0.042	0.042	0.036	0.41	0.77	0.042	0.041	0.19	0.25
$\lambda_{4[12]}$	0.009	0.010	0.007	2.93	4.53	0.009	0.009	0.86	1.15
μ_2	0.147	0.150	0.108	1.39	0.86	0.147	0.135	0.45	0.36
$\xi_{2[12]}$	0.033	0.034	0.019	2.68	1.86	0.032	0.028	0.70	0.63
ρ_{12}	0.222	0.204	0.181	0.98	1.49	0.217	0.267	0.39	0.50
μ_3	0.308	0.413	0.138	7.56	1.55	0.388	0.247	3.72	0.77
$\xi_{3[12]}$	0.068	0.101	0.019	12.5	2.96	0.083	0.046	3.33	1.11
μ_4	3.227	2.517	0.185	25.4	2.50	2.485	0.551	19.7	1.40
$\xi_{4[12]}$	0.717	0.738	0.019	43.3	4.66	0.452	0.091	16.0	1.86