Multivariate Extension of L-Moments via L-Comoments, and Applications

Peng Xiao and Robert Serfling[1]

East Carolina University and University of Texas at Dallas

Joint Statistical Meetings 2008
Denver

[1] www.utdallas.edu/∼serfling

FOCAL TOPICS OF THIS TALK

L-Comoment Matrices: A *Pairwise* Multivariate Approach Toward Disperson, Correlation, Skewness, Kurtosis, etc.

L-Correlation: *First Order Correlation Analysis*

Asymptotic Behavior: *Consistency* and *Limit Distributions*

Trimmed Versions for More Robustness and Lower Moment Assumptions

Applications in *Multivariate Regional Frequency Analysis* and *Financial Portfolio Analysis*

Moment Assumptions in Multivariate Analysis

- Multivariate statistical analysis typically uses moment assumptions of *second order* and higher.
- Increasing interest, however, in *heavy tailed* distributions.
- Can we describe multivariate *dispersion*, *skewness*, and *kurtosis*, and have *correlational analysis*, under just *first order* moment assumptions?
- Goals in both *parametric* and *nonparametric* settings.

A Partial Solution: Pairwise Treatment of Variables

- We may not think properly about high dimensional descriptive measures (curse of dimensionality).
- But we understand *bivariate* distributions very well.
- Pairwise treatment is used already for *covariance and correlation matrices* and *pairwise scatterplots*.
- Can "pairwise" be a *solution* instead of an *apology*?
- Should we develop a *comprehensive* pairwise approach that includes handling of *skewness* and *kurtosis*?

First, A Look at Classical Central Comoments

- d-vector $\mathbf{X} = (X^{(1)}, \ldots, X^{(d)})'$
- marginal means μ_i and central moments $\mu_k^{(i)}$, $k \geq 2$
- *covariance* of $X^{(i)}$ and $X^{(j)}$

$$\mathrm{Cov}\{X^{(i)}, X^{(j)}\} = \mathrm{Cov}\{X^{(i)} - \mu_i, X^{(j)} - \mu_j\}$$

- *coskewness of* $X^{(i)}$ *with respect to* $X^{(j)}$

$$\mathrm{Cov}\{X^{(i)} - \mu_i, (X^{(j)} - \mu_j)^2\}$$

- *cokurtosis of* $X^{(i)}$ *with respect to* $X^{(j)}$

$$\mathrm{Cov}\{X^{(i)} - \mu_i, (X^{(j)} - \mu_j)^3\}$$

Central Comoment Matrices

- *kth order central comoment matrix*:

$$(\xi_{k\,[ij]}) = \left(\mathrm{Cov}\{X^{(i)} - \mu_i, (X^{(j)} - \mu_j)^{k-1}\}\right)$$

- Uses the *covariance operator* element-wise.
- For order $k \geq 3$, *asymmetric*, which is *good* because more information is provided.
- Symmetry for $k = 2$ (the classical covariance matrix) is an unfortunate exception, merely an artifact of the definition.

Central Comoments Not New

- *Central comoments are not new*:
  Rubinstein, M. E. (1973). The fundamental theorem of parameter-preference security valuation. *Journal of Financial and Quantitative Analysis* **8** 61–69.

- *But a new key question*:

  Can a useful notion of kth comoment be defined under just **first moment** *assumptions?*

  For example, can we have a correlation analysis with just first moments finite?

Next, A Look at Representations for L-Moments

- *As an L-functional*

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

using shifted Legendre orthogonal polynomials: $P^*_k(u) = \sum_{j=0}^k p^*_{k,j} u^j$, $0 \le u \le 1$, with $p^*_{k,j} = (-1)^{k-j}\binom{k}{j}\binom{k+j}{j}$.

- *As a covariance*

$$\boxed{\lambda_k = \mathrm{Cov}\{X, P^*_{k-1}(F(X))\},\ \ k \ge 2}.$$

Put These Together to Define L-Comoments

- Combine the covariance representation for L-moments,

$$\lambda_k = \mathrm{Cov}\{X, P^*_{k-1}(F(X))\},\ \ k \geq 2,$$

with the definition of kth order central comoment

$$\mathrm{Cov}\{X^{(i)} - \mu_i, (X^{(j)} - \mu_j)^{k-1}\},\ \ k \geq 2.$$

- Thus define the *kth order L-comoment*:

$$\boxed{\lambda_{k\,[ij]} = \mathrm{Cov}(X^{(i)}, P^*_{k-1}(F_j(X^{(j)}))),\ \ k \geq 2},$$

with F_j the cdf of $X^{(j)}$. Note that the L-comoments are asymmetric for *all* $k \geq 2$.

A Representation in Terms of Concomitants

- For a bivariate sample $(X_m^{(i)}, X_m^{(j)}), 1 \leq m \leq n$, with ordered $X^{(j)}$-values $X_{1:n}^{(j)} \leq X_{2:n}^{(j)} \leq \cdots \leq X_{n:n}^{(j)}$, the *concomitant* of $X_{r:n}^{(j)}$ is the $X_m^{(i)}$ paired with it, say $X_{[r:n]}^{(ij)}$ (David and Nagaraja, 2003).
- Then

$$\lambda_{k\,[ij]} = k^{-1} \sum_{r=0}^{k-1} (-1)^r \binom{k-1}{r} E(X_{[k-r:k]}^{(ij)}),$$

a formula for λ_k with $E(X_{k-r:k})$ replaced by $E(X_{[k-r:k]}^{(ij)})$.

Another Representation, Yielding a Sample Version

- We also have

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

  a formula for λ_k with $E(X_{r:n})$ replaced by $E(X_{[r:n]}^{(ij)})$.

- Replacing $E(X_{[r:n]}^{(ij)})$ by $X_{[r:n]}^{(ij)}$ yields an unbiased estimator for $\lambda_{k\,[ij]}$ based on a sample of size n.

- The 2nd L-comoment is the *Gini covariance* studied by Schectman and Yitzhaki (1987), Yitzhaki and Olkin (1991), and Olkin and Yitzhaki (1992).
- The *L-correlation* of $X^{(i)}$ with respect to $X^{(j)}$ is defined by scaling the corresponding 2nd L-comoment:

$$\rho_{[ij]} = \frac{\lambda_{2\,[ij]}}{\lambda_2^{(i)}}.$$

- This is defined under just a 1st moment assumption on $X^{(i)}$. Note that $\rho_{[ij]}$ and $\rho_{[ji]}$ need not be equal.

The Usual Bounds Are Satisfied

- **Proposition** The L-correlation satisfies

$$|\rho_{[ij]}| \leq 1,$$

with equality if $X^{(i)}$ is a strictly monotone function of $X^{(j)}$ (a necessary condition in the case of continuous distributions).

A Lemma on 2nd L-Moments of Sums

- We ask: *What is the L-moment analogue of the decomposition of the variance of a sum in terms of variances of the summands?*

A Lemma on 2nd L-Moments of Sums

- We ask: *What is the L-moment analogue of the decomposition of the variance of a sum in terms of variances of the summands?*
- Answer: *The 2nd L-moment of a sum is a weighted sum of the 2nd L-moments of the summands, with the coefficients given by L-correlations.*

The Proof is Immediate

For univariate $Y_1, \ldots, Y_n$ and $S_n = Y_1 + \cdots + Y_n$,

$$\begin{aligned}\lambda_2(S_n) &= 2\,\mathrm{Cov}(S_n, F_{S_n}(S_n)) = 2\sum_{i=1}^{n} \mathrm{Cov}(Y_i, F_{S_n}(S_n)) \\ &= \sum_{i=1}^{n} \lambda_{2[12]}(Y_i, S_n) = \boxed{\textstyle\sum_{i=1}^{n} \rho_{[Y_i, S_n]}\, \lambda_2(Y_i)}\end{aligned}$$

(*without any dependence restrictions on the summands*).

When Does L-Correlation Agree with Pearson Correlation?

- **Proposition** Assume (i) $X^{(1)}$ has *linear regression* on $X^{(2)}$: $E(X^{(1)} \mid X^{(2)}) = a + bX^{(2)}$, and (ii) $X^{(1)}$ and $X^{(2)}$ are *affinely equivalent in distribution*: $X^{(2)} \stackrel{d}{=} \theta + \eta X^{(1)}$. Then, under 2nd moment assumptions,
$$\rho_{[12]} = \rho_{12} \quad \text{(the usual Pearson correlation).}$$
- Thus, under (i) and (ii), L-correlation coherently extends Pearson correlation down to 1st moment assumptions.

Illustration: A Multivariate Pareto Distribution

- The Type II multivariate Pareto distribution of Arnold (1983) is

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

for $x^{(i)} > \theta_i$ and $\sigma_i > 0$, $1 \le i \le d$, and $\alpha > 0$.

- The kth marginal moments are finite if $k < \alpha$.
- Many typical applications involve heavy-tailed modeling, with α in the range 1 to 2 for quite diverse data sets

Correlation Analysis for This Model

- This model satisfies conditions (i) and (ii) for agreement between L-correlation and Pearson correlation when both are defined.
- Both correlations have the formula α^{-1}, for $\alpha > 2$.
- The L-correlation also has this same formula for $1 < \alpha \leq 2$.
- We thus obtain for this model an extended correlation analysis, from the range $\alpha > 2$ to the range $\alpha > 1$.
- The maximal value 1/2 possible for the correlation under the restriction $\alpha > 2$ becomes 1 under $\alpha > 1$ and this value is approached as $\alpha \downarrow 1$.

Results on Asymptotic Behavior are Straightforward

- By analogy with handling of univariate sample L-moments.
- Using a representation of the kth order sample L-comoment as a U-statistic.
- Or using results on asymptotic normality of linear functions of concomitants (Yang, 1981).

But Under Just First Moment Assumptions?

- U-statistics are *consisent* under just a first moment assumption on the kernel.
- *Asymptotic normality* of U-statistics requires, however, somewhat higher moment assumptions.
- Asymptotic normality of *L-statistics* with weight functions nonzero on pth quantiles for p near 0 or 1 also requires somewhat higher moment assumptions.

More Robustness and Lower Moment Assumptions

- A kth *trimmed L-comoment* uses concomitants of the k order statistics remaining after trimming the t_1 smallest and t_2 largest in a conceptual sample of size $k + t_1 + t_2$ and thus replaces $E(X^{(ij)}_{[k-r:k]})$ by $E(X^{(ij)}_{[k+t_1-r:k+t_1+t_2]})$:

$$\lambda_k^{(t_1,t_2)} = k^{-1} \sum_{r=0}^{k-1} (-1)^r \binom{k-1}{r} E(X^{(ij)}_{[k+t_1-r:k+t_1+t_2]}),$$

- Analogous to the trimmed L-moments of Elamir and Seheult (2003), this permits moment assumptions lower than 1st order and eliminates the influence of the most extreme sample points.

Univariate Regional Frequency Analysis

- Regional frequency analysis estimates extreme hydrological events (droughts, flood peaks or volumes, rain storms, etc.) at sites with little or no data.
- Hosking and Wallis (1997) [*Regional Frequency Analysis: An Approach Based on L-Moments*, Cambridge University Press] group sites into homogeneous regions with respect to a chosen variable and fit regional distributions.
- This involves measures of *discordancy of a site*, *heterogeneity of a region*, and *goodness-of-fit of a regional distribution*, via a method of L-moments based on the 1st, 2nd, 3rd, and 4th L-moments.

Multivariate Regional Frequency Analysis

- *Problem*: Hydrological events involve many variables, and a thorough understanding requires treating them jointly.
- *Solution*: Chebana and Ouarda (2007) [Multivariate L-moment homogeneity test, *Water Resources Research* **43**] extend the univariate approach of Hosking and Wallis (1997) to the multivariate case.
- The role of 2nd, 3rd, and 4th L-moments is played by *matrix norms* of the correspondng L-comoment matrices of Serfling and Xiao (2007).

Portfolio Optimization in Finance

- A central role has long been played by the capital asset pricing model (CAPM), initially involving just 1st and 2nd moments but recently higher moments also.
- Skewness measures concern evaluation of downside risk and asymmetric volatility of a portfolio, while spread and kurtosis measures concern volatility.
- Increasing interest in heavy tailed distributions in modeling stock returns raises serious concern regarding higher moment assumptions and lack of stability and robustness associated with higher-order central moments and comoments.

L-Moment Approach

- For the *marginal distributions* of jointly distributed heavy tailed variables in risk analysis, univariate L-moments have already been applied in Hosking, Bonti, and Siegel (2000) [Beyond the lognormal. *Risk* **13** 59–62].

L-Comoment Approach

- A *multivariate* nonparametric portfolio selection approach based on the L-comoments is developed by Jurczenko, Maillet, and Merlin (2008) [Efficient Frontier for Robust Higher-Moment Portfolio Selection. Preprint].

L-Comoment Approach

- A *multivariate* nonparametric portfolio selection approach based on the L-comoments is developed by Jurczenko, Maillet, and Merlin (2008) [Efficient Frontier for Robust Higher-Moment Portfolio Selection. Preprint].
- For a collection of securities, the weights for an optimal portfolio are found that solve the multi-objective problem of simultaneously maximizing the weighted averages of the 1st and 3rd order L-comoments, and minimizing those of the 2nd and 4th order L-comoments, of the securities' returns with the portfolio return.

Acknowledgment

Support of the second author by grants from NSF, NSA, and the Society of Actuaries is greatly appreciated.

Reference

- Serfling, R., and Xiao, P. (2007). A contribution to multivariate L-moments: L-comoment matrices. *Journal of Multivariate Analysis* **98** 1765–1781.

- Links to this paper, to the present talk (pdf), and to related software programs, are found at `www.utdallas.edu/~serfling`.