Inference for models defined through L-moment conditions

Michel Broniatowski, Alexis Decurninge

LSTA-Université Pierre et Marie Curie

February 2016

Contents

For univariate distributions, L-moments are expressed as the expectation of a particular linear combination of order statistics. $X_1, ..., X_r$ iid copies of X with $\mathbb{E}(|X|) < \infty$. The r-th L-moment is defined by

$$\lambda_r = \frac{1}{r} \sum_{k=0}^{r-1} (-1)^k \binom{r-1}{k} \mathbb{E}[X_{r-k:r}] \tag{1}$$

where $X_{1:r} \leq ... \leq X_{r:r}$ denotes the order statistics. The four first L-moment can be considered as a measure of location, dispersion, skewness and kurtosis. Indeed $\lambda_1 = \mathbb{E}(X)$, $\lambda_2 = (1/2)\,\mathbb{E}(|X - Y|)$ with Y an independent copy of . λ_3 : the expected distance between the mean of the extreme terms and the median one in a sample of three i.i.d. replications of X, and λ_4 :the expected distance between the extreme terms of a sample of four replicates of X with respect to a multiple of the distance between the two central terms.

L-moments constitute a robust alternative to traditional moments. Applications dealing with heavy-tailed distributions. Hosking:"The main advantage of L-moments over conventional moments is that L-moments, being linear functions of the data, suffer less from the effect 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. Also as seen through (1) the L-moments are determined by the expectation of extreme order statistics, and vice versa". This motivates their success for the inference in models pertaining to the tail behavior of random phenomenons.

Example

The family of all the distributions of a r.v. X whose second, third and fourth L-moments verify :

$$\left\{\begin{array}{l} \lambda_2 = \sigma(1-2^{-1/\nu})\Gamma(1+1/\nu) \\ \lambda_3 = \lambda_2[3-2\frac{1-3^{-1/\nu}}{1-2^{-1/\nu}})] \\ \lambda_4 = \lambda_2[6+\frac{5(1-4^{-1/\nu})-10(1-3^{-1/\nu})}{1-2^{-1/\nu}})] \end{array}\right. \quad (2)$$

for any $\sigma > 0, \nu > 0$. These distributions share their first L-moments of order 2, 3 and 4 with those of a Weibull distribution with scale and shape parameter σ and ν. Shift invariance. the r.v. Y shares the same L-moments λ_r with those of X but for $r = 1$: a neighborhood of the continuum of all Weibull distributions on $[a, \infty)$ or on $(-\infty, a]$ when a belongs to $\mathbb{R}$. Hence this model aims at describing a shape constraint on the tail of the distribution of the data, independently of its location.

Example

The model which is the space of the distributions whose second, third and fourth L-moments verify :

$$\begin{cases} \lambda_2 = \frac{\sigma}{(1+\nu)(2+\nu)} \\ \lambda_3 = \lambda_2 \frac{1-\nu}{3+\nu} \\ \lambda_4 = \lambda_2 \frac{(1-\nu)(2-\nu)}{(3+\nu)(4+\nu)} \end{cases} \tag{3}$$

for any $\sigma > 0, \nu \in \mathbb{R}$. These distributions share their first L-moments with those of a generalized Pareto distribution with scale and shape parameter σ and ν. A neighborhood of the whole continuum of Pareto distributions on $[a, \infty)$ or on $(-\infty, a]$ when a belongs to $\mathbb{R}$.

Open and interesting questions: how "large are those neighbourhoods", etc
For example

Proposition

Suppose X is positive with finite expectation (its quantile function is denoted by Q). Let $u_0 > 1/2$.

$$\mathbb{P}[X > Q(u_0)] \leq \frac{\lambda_2}{Q(u_2) - Q(u_1)}$$

with

$$u_1 = \frac{1 - \sqrt{8u_0 - 7}}{2} \qquad u_2 = \frac{1 + \sqrt{8u_0 - 7}}{2}$$

L-moments

Definition and characterizations

The r-th L-moment λ_r is defined as a particular expectation of L-statistics

$$\lambda_r = \frac{1}{r}\sum_{k=0}^{r-1}(-1)^k\binom{r-1}{k}\mathbb{E}[X_{r-k:r}] \tag{4}$$

The first four L-moments are

$$\begin{aligned}
\lambda_1 &= \mathbb{E}[X] \\
\lambda_2 &= \tfrac{1}{2}\mathbb{E}[X_{2:2} - X_{1:2}] \\
\lambda_3 &= \tfrac{1}{3}\mathbb{E}[X_{3:3} - 2X_{2:3} + X_{1:3}] \\
\lambda_4 &= \tfrac{1}{4}\mathbb{E}[X_{4:4} - 3X_{3:4} + 3X_{2:4} - X_{1:4}].
\end{aligned}$$

The expectations of the extreme order statistics characterize a distribution: if $\mathbb{E}\left(|X|\right)$ is finite, either of the sets $\{\mathbb{E}\left(X_{1:n}\right), n = 1, ..\}$ or $\{\mathbb{E}\left(X_{n:n}\right), n = 1, ..\}$ characterize the distribution of X.;hence L-moments characterize the distribution of X.

Writing the L-moments of a distribution F as an inner product of the corresponding quantile function with a specific complete orthonormal system of polynomials in $L^2(0,1)$. The shifted Legendre polynomials define such a system of functions.

Definition

The shifted Legendre polynomial of order r is

$$L_r(t) = \sum_{k=0}^{r} (-1)^k \binom{r}{k}^2 t^{r-k}(1-t)^k = \sum_{k=0}^{r} (-1)^{r-k} \binom{r}{k}\binom{r+k}{k} t^k. \tag{5}$$

Let us define for $r \geq 1$, K_r as the integrated shifted Legendre polynomials

$$K_r(t) = \int_0^t L_{r-1}(u)du = -t(1-t)\frac{J_{r-2}^{(1,1)}(2t-1)}{r-1} \tag{6}$$

with $J_{r-2}^{(1,1)}$ the corresponding Jacobi polynomial

All L-moments λ_r but λ_1 are shift invariant, hence independent upon λ_1.
If F is continuous, the expectation of the j-th order statistics $X_{j:r}$ is

$$\mathbb{E}[X_{j:r}] = \frac{r!}{(j-1)!(r-j)!} \int_{\mathbb{R}} xF(x)^{j-1}(1 - F(x))^{r-j}\mathbf{F}(dx). \tag{7}$$

We can state the following result.
If $r \geq 2$ and $\int_{\mathbb{R}} |x|\, dF(x) < +\infty$, then

$$\lambda_r = \int_0^1 F^{-1}(t)dK_r(t) = -\int_0^1 K_r(t)\mathbf{F}^{-1}(dt). \tag{8}$$

Estimation of L-moments

Let $x_1, \ldots, x_n$ be iid realizations of a random variable X with distribution F and L-moments λ_r. Define F_n the empirical cdf of the sample and l_r the corresponding plug-in estimator of λ_r,

$$l_r = \int_0^1 F_n^{-1}(t) L_{r-1}(t) dt. \tag{9}$$

This estimator of λ_r is biased; the unbiased estimators of L-moments are the following U-statistics

$$l_r^{(u)} = \frac{1}{\binom{n}{r}} \sum_{1 \le i_1 < \cdots < i_r \le n} \frac{1}{r} \sum_{k=0}^{r-1} (-1)^k \binom{r-1}{k} x_{i_{r-k}:n}.$$

These two estimators l_r and $l_r^{(u)}$ of the L-moment λ_r have the same asymptotic properties.

Models

Constraints on moments

Let θ in Θ , an open subset of $\mathbb{R}^d$ and let $g : (x, \theta) \in \mathbb{R} \times \Theta \to \mathbb{R}^l$ be a l-valued function, each component of which is parametrized by $\theta \in \Theta \subset \mathbb{R}^d$. Define

$$M_\theta := \left\{ \mathbf{F} \text{ s.t. } \int_{\mathbb{R}} g(x, \theta)\mathbf{F}(dx) = 0 \right\}$$

and the *semi parametric* model defined by moment conditions is the collection of probability measures in

$$\mathcal{M} := \bigcup_{\theta \in \Theta} M_\theta. \tag{10}$$

These semiparametric models are defined by l conditions pertaining to l moments of the distributions and are widely used in applied statistics. When $d > l$: Hansen (**GMM**); Owen (**empirical likelihood approach)**. Newey and Smith or B and Keziou (**empirical divergence approach)**. *linearity with respect to the cumulative distribution function* (cdf) which brings a *dual formulation of the optimization problem* for estimation.

Constraints on L-moments

Similarly as for models defined by (10), we can introduce *semiparametric linear quantile* (SPLQ) models through

$$\bigcup_{\theta\in\Theta} L_\theta := \bigcup_{\theta\in\Theta} \left\{ \mathbf{F} \text{ s.t. } \int_0^1 F^{-1}(u) k(u,\theta) du = f(\theta) \right\} \qquad (11)$$

where $\Theta \subset \mathbb{R}^d$, $k : (u,\theta) \in [0;1] \times \Theta \to \mathbb{R}^l$ and $f : \Theta \to \mathbb{R}^l$. Here F^{-1} denotes the generalized inverse function of F, the distribution function of the measure $\mathbf{F}$. A specific choice for $k(u,\theta)$ produces constraints on the $L-$moments of $\mathbf{F}$.

Example

$$\begin{aligned} \lambda_r &= \int_0^1 F^{-1}(t) dK_r(t) \quad 2 \le r \le l \\ \lambda_r &= f_r(\theta). \end{aligned}$$

Typically models defined by a finite number of constraints on functions of the moments of the order statistics.

Models defined by L-moments conditions

We consider models defined by l constraints on their first L-moments

$$-\mathbb{E}\left[\frac{1}{r}\sum_{k=0}^{r-1}(-1)^k\binom{r-1}{k}X_{k:r}\right] = f_r(\theta) \quad 2 \leq r \leq l \tag{12}$$

where Θ is some open set in $\mathbb{R}^d$ and $f_j : \Theta \to \mathbb{R}$ are some given functions defined on Θ.

Semi Parametric Linear Quantile Models, with $(u,\theta) \mapsto k(u,\theta)$ independent on θ, defined by

$$k(u,\theta) = -L(u) := -\begin{pmatrix} L_2(u) \\ \vdots \\ L_l(u) \end{pmatrix} \tag{13}$$

The SPLQ model (11) may be written as

$$\mathcal{L} := \bigcup_{\theta\in\Theta} L_\theta = \bigcup_{\theta\in\Theta}\left\{\mathbf{F} \text{ s.t. } \int_0^1 L(u)F^{-1}(u)du = -f(\theta)\right\}. \tag{14}$$

We may write equation (12) for $r \geq 2$ as follows, making use of the integrated shifted Legendre polynomials K_r in place of L_r.

$$-\mathbb{E}\left[\frac{1}{r}\sum_{k=0}^{r-1}(-1)^k\binom{r-1}{k}X_{k:n}\right] = \int_0^1 K_r(u)\mathbf{F}^{-1}(du) = f_r(\theta). \quad (15)$$

Example

Weibull type: we define k and f by

$$k(u,\theta) = -\begin{pmatrix} L_2(u) \\ L_3(u) \\ L_4(u) \end{pmatrix}$$

and

$$f(\theta) = \begin{pmatrix} f_2(\theta) \\ f_3(\theta) \\ f_4(\theta) \end{pmatrix} = \begin{pmatrix} \sigma(1-2^{-1/\nu})\Gamma(1+1/\nu) \\ f_2(\theta)[3-2\frac{1-3^{-1/\nu}}{1-2^{-1/\nu}})] \\ f_2(\theta)[6+\frac{5(1-4^{-1/\nu})-10(1-3^{-1/\nu})}{1-2^{-1/\nu}})] \end{pmatrix}$$

where $\theta = (\sigma,\nu) \in \mathbb{R}_+^* \times \mathbb{R}_+^*$ and $u \in [0;1]$.

Example

Pareto type: we define k and f by

$$k(u,\theta) = -\begin{pmatrix} L_2(u) \\ L_3(u) \\ L_4(u) \end{pmatrix}$$

and

$$f(\theta) = \begin{pmatrix} f_2(\theta) \\ f_3(\theta) \\ f_4(\theta) \end{pmatrix} = \begin{pmatrix} \frac{\sigma}{(1+\nu)(2+\nu)} \\ f_2(\theta)\frac{1-\nu}{3+\nu} \\ f_2(\theta)\frac{(1-\nu)(2-\nu)}{(3+\nu)(4+\nu)} \end{pmatrix}$$

where $\theta = (\sigma,\nu) \in \mathbb{R}_+^* \times \mathbb{R}$ and $u \in [0;1]$.

Extension to models defined by order statistics conditions

The moments of order statistics given by equation satisfy

$$\mathbb{E}[X_{j:r}] = \int_0^1 P_{j:r}(u) F^{-1}(u) du$$

where

$$P_{j:r}(u) = \frac{r!}{(j-1)!(r-j)!} u^{j-1}(1-u)^{r-j}.$$

Any expectation of the moment of some L-statistics writes

$$-\sum_{i=1}^{r} a_j \mathbb{E}[X_{j:r}] = \int_0^1 P_a(u) F^{-1}(u) du, \; P_a(u) = -\sum_{i=1}^{r} a_j P_{j:r}(u).$$

These models are SPLQ with

$$\mathcal{L} := \bigcup_{\theta} L_\theta = \bigcup_{\theta} \left\{ F \text{ s.t. } \int_0^1 P(u) F^{-1}(u) du = -f(\theta) \right\} \tag{16}$$

where $P : u \in [0;1] \mapsto P(u) \in \mathbb{R}^l$ is an array of l polynomials.

φ-divergences

Let $\varphi : \mathbb{R} \to [0, +\infty]$ be a strictly convex function with $\varphi(1) = 0$ such that $\text{dom}(\varphi) = \{x \in \mathbb{R} | \varphi(x) < \infty\} := (a_\varphi, b_\varphi)$ with $a_\varphi < 1 < b_\varphi$. If $F << G$ two σ-finite measures of $(\mathbb{R}, B(\mathbb{R}))$ such that

$$D_\varphi(G, F) = \int_{\mathbb{R}} \varphi\left(\frac{dG}{dF}(x)\right) dF(x) \qquad (17)$$

$F = G$ iff $D_\varphi(F, G) = 0$, when φ srtict convex.

Example

The class of power divergences parametrized by $\gamma \geq 0$ is defined through the functions

$$x \mapsto \varphi_\gamma(x) = \frac{x^\gamma - \gamma x + \gamma - 1}{\gamma(\gamma - 1)}.$$

The domain of φ_γ depends on γ. The Kullback-Leibler divergence is associated to $x > 0 \mapsto \varphi_1(x) = x \log(x) - x + 1$, the modified Kullback-Leibler (KL_m) divergence to $x > 0 \mapsto \varphi_0(x) = -\log(x) + x - 1$, the χ^2-divergence to $x \in \mathbb{R} \mapsto \varphi_2(x) = 1/2(x-1)^2$, etc.

Estimates with L-moments constraints

Minimum of φ-divergences for p.m's under L-moment constraints

A plain approach to inference on θ : mimick the empirical minimum divergence one, substituting the linear constraints with respect to the distribution

$$\int_{\mathbb{R}} g(x,\theta)\mathbf{F}(dx) = \mathbb{E}[g(X,\theta)] = 0$$

by the corresponding linear constraints with respect to the quantile measure

$$L_\theta^{(0)}(\mathbf{F}) = \{\mathbf{G} \text{ s.t. } \mathbf{G} \ll \mathbf{F}, \int_0^1 K(u)\mathbf{G}^{-1}(du) = f(\theta)\}$$

For any parameter $\theta \in \Theta$, the distance between $\mathbf{F}$ and the submodel $L_\theta^{(0)}(\mathbf{F})$ is defined by

$$D_\varphi(L_\theta^{(0)}(\mathbf{F}), \mathbf{F}) = \inf_{\mathbf{G} \in L_\theta^{(0)}(\mathbf{F})} D_\varphi(\mathbf{G}, \mathbf{F}),$$

Hence minimize the divergence between all probability measures satisfying the constraint and the empirical measure $\mathbf{F}_n$ pertaining to the data set.
A **Plug-in estimator** is

$$D_\varphi(L_\theta^{(0)}(\mathbf{F}_n), \mathbf{F}_n) = \inf_{\mathbf{G} \in L_\theta^{(0)}(\mathbf{F}_n)} D_\varphi(\mathbf{G}, \mathbf{F}_n).$$

A natural estimator for θ may be defined by

$$\begin{aligned} \hat{\theta}_n^{(0)} &= \arg\inf_{\theta \in \Theta} D_\varphi(L_\theta^{(0)}(\mathbf{F}_n), \mathbf{F}_n) \\ &= \arg\inf_{\theta \in \Theta} \inf_{\mathbf{G} \in L_\theta^{(0)}(F_n)} \frac{1}{n} \sum_{i=1}^{n} \varphi(n\mathbf{G}(x_i)). \end{aligned}$$

Existence of this estimator may not hold. $L_\theta^{(0)}(\mathbf{F}_n)$ may be void: its elements are multinomial distributions $\sum_{i=1}^n w_i \delta_{x_i}$ whose weights are solutions of a family of $l-1$ polynomial algebraic equation of degree l (with n unknowns $w_1, ..., w_n$).

$$\lambda_r = -\int_0^1 K_r(t)\mathbf{G}^{-1}(dt) \quad ; 2 < r \leq l. \tag{18}$$

takes the form

$$\sum_{i=1}^n K_r\left(\sum_{a=1}^i w_a\right)(x_{i+1:n} - x_{i:n}) = -f_r(\theta);\ 2 < r \leq l.$$

By (18) the quantile function plays a similar role as the distribution function in the classical moment equations. We will then change the functional to be minimized in order to be able to use duality for the optimization step. Divergences between quantile measures.

Minimum of φ-divergences for quantile measures, primal problem

For any θ in Θ the submodel L_θ^n is Let N be the class of all $\sigma-$finite positive measures on $\mathbb{R}$. Making use of equation (15) define

$$L_\theta^n := \left\{ \begin{array}{c} \mathbf{G}^{-1} \in N \text{ s.t. } \mathbf{G}^{-1} \ll \mathbf{F}_n^{-1} \text{ and } -\int_0^1 L(u) G^{-1}(u) du \\ = \int_0^1 K(u) \mathbf{G}^{-1}(du) = f(\theta) \end{array} \right\}$$

the family of all measures $\mathbf{G}^{-1}$ with support included in $\{1/n, 2/n, .., (n-1)/n\}$ which satisfy the $l-1$ constraints pertaining to the L-moments.

A natural proposal for an estimation procedure in the SPLQ model is then to consider the minimum of a φ-divergence between quantile measures through

$$\hat{\theta}_n = \arg \inf_{\theta \in \Theta} \inf_{\mathbf{G}^{-1} \in L_\theta^n} \int_0^1 \varphi \left(\frac{d\mathbf{G}^{-1}}{d\mathbf{F}_n^{-1}}(u) \right) \mathbf{F}_n^{-1}(du)$$

$$= \arg \inf_{\theta \in \Theta} \inf_{\Delta} \sum_{i=1}^{n-1} \varphi \left(\frac{y_{i+1} - y_i}{x_{i+1:n} - x_{i:n}} \right) (x_{i+1:n} - x_{i:n})$$

$$\Delta := \left\{ (y_{i+1} - y_i \geq 0) \, , 1 \leq i \leq n-1, \sum_{i=1}^{n-1} K(i/n)(y_{i+1} - y_i) = f(\theta) \right\}.$$

Both the constraint and the divergence criterion are expressed in function of $\mathbf{G}^{-1}$ and the constraint is linear with respect to this measure. This allows to use classical duality results in order to efficiently compute the estimator $\hat{\theta}_n$.Condition $(y_{i+1} - y_i \geq 0) \, 1 \leq i \leq n-1$ may be relaxed when $\mathbf{G}^{-1}$ is a $\sigma-$ finite signed measure, nonwhithstanding notation (!) .

Dual representations of the divergence under L-moment constraints

The minimization of φ-divergences under linear equality constraint is performed using Fenchel-Legendre duality. It transforms the constrained problems into an unconstrained one in the space of Lagrangian parameters. Let ψ denote the Fenchel-Legendre transform of φ, namely, for any $t \in \mathbb{R}$

$$\psi(t) := \sup_{x \in \mathbb{R}} \{tx - \varphi(x)\}.$$

Let $\theta \in \Theta$ and F be fixed.

Theorem

If there exists some $\mathbf{G}^{-1}$ *in* $L_\theta(\mathbf{F}^{-1})$ *such that* $a_\varphi < d\mathbf{G}^{-1}/d\mathbf{F}^{-1} < b_\varphi$ $\mathbf{F}^{-1}$*-a.s. then*

$$\inf_{\mathbf{G}^{-1} \in L_\theta(\mathbf{F}^{-1})} \int_0^1 \varphi\left(\frac{d\mathbf{G}^{-1}}{d\mathbf{F}^{-1}}\right) d\mathbf{F}^{-1} = \sup_{\xi \in \mathbb{R}^l} \xi^T f(\theta) - \int_0^1 \psi(\xi^T K(u))\mathbf{F}^{-1}(du). \tag{19}$$

Moreover, if ψ *is differentiable and if there exists a solution* ξ^* *of the dual problem which is an interior point of*

$$\left\{\xi \in \mathbb{R}^l \text{ s.t. } \int_{\mathbb{R}} \psi(\xi^T K(u))\mathbf{F}^{-1}(du) < \infty\right\},$$

then ξ^* *is the unique maximum in (19) and*

$$\int \psi'\left((\xi^*)^T K(u)\right) K(u)\mathbf{F}^{-1}(du) = f(\theta).$$

The estimator

Dual Problem for all θ

$$\inf_{\mathbf{G}^{-1}\in L_\theta(\mathbf{F}_n^{-1})} \int_0^1 \varphi\left(\frac{d\mathbf{G}^{-1}}{d\mathbf{F}_n^{-1}}\right) d\mathbf{F}_n^{-1} = \sup_{\xi\in\mathbb{R}^l} \langle \xi, f(\theta)\rangle - \int_0^1 \psi(\langle \xi, K(u)\rangle)\mathbf{F}_n^{-1}(du)$$

Hence

$$\theta_n := \arg\inf_{\theta} \sup_{\xi\in\mathbb{R}^l} \xi^T f(\theta) - \int_0^1 \psi(\xi^T K(u))\mathbf{F}_n^{-1}(du).$$

Example

The χ^2-divergence $\varphi(x) = \frac{(x-1)^2}{2}$, then $\psi(t) = \frac{1}{2}t^2 + t$ and for each θ

$$\xi_1^* = \Omega_n^{-1}\left(f(\theta) - \int K(F_n(x))d\lambda\right)$$

with

$$\Omega_n = \int K(F_n(x))K(F_n(x))^T d\lambda$$

Hence the estimator

$$\hat{\theta}_n = \arg\inf_{\theta\in\Theta} d_n\Omega_n^{-1}d_n$$

with

$$d_n := f(\theta) - \int K(F_n(x))d\lambda$$

shares similarities with the GMM estimator.
This divergence should thus be favored for its fast implementation.

Asymptotic properties of the estimators under L-moment conditions

Theorem

Let $x_1, \dots, x_n$ be an observed sample drawn iid from a distribution F_0 with finite expectation. Let us suppose that

- *there exists θ_0 such that $F_0 \in L_{\theta_0}$, θ_0 is the unique solution of the equation $f(\theta) = f(\theta_0)$*
- *f is continuous and $\Theta \subset \mathbb{R}^d$ is compact*
- *the matrix $\Omega_0 = \int K(F_0(x))K(F_0(x))^T dx$ is non singular.*

Then

$$\hat{\theta}_n \longrightarrow \theta_0 \text{ in probability as } n \to \infty.$$

Theorem

Let $x_1, ..., x_n$ be an observed sample drawn iid from a distribution F_0 with finite variance. Assume regularity as above and

- $\theta_0 \in int(\Theta)$
- $J_0 = J_f(\theta_0)$ *the Jacobian of* f *with respect to* θ *in* θ_0 *has full rank*
- f *is continuously differentiable in a neighborhood of* θ_0

Then

$$\sqrt{n}\begin{pmatrix} \hat{\theta}_n - \theta_0 \\ \hat{\xi}_n \end{pmatrix} \to_d \mathcal{N}\left(0, \begin{pmatrix} H\Sigma H^T & 0 \\ 0 & P\Sigma P^T \end{pmatrix}\right)$$

with

$$\begin{aligned} \Sigma &:= \int\int \left[F(\min(x,y) - F(x)F(y))\, K'(F(x)K'(F(y)^T dxdy\right] \\ H &:= MJ_0^T\Omega^{-1} \\ M &:= \left(J_0^T\Omega^{-1}J_0\right)^{-1} \\ P &:= \Omega^{-1} - \Omega^{-1}J_0MJ_0^T\Omega^{-1}. \end{aligned}$$

The estimator of the divergence from **F** onto the model, namely

$$2n\left[\xi_n^T f(\theta_n) - \int \psi\left(\xi_n^T K\left(F_n(t)\right)\right) dt\right]$$

does not converge to a χ^2 distribution as in moment condition models. However

$$S_n := n\xi_n^T\left(P_n\Sigma_n P_n^T\right) - \xi_n \rightarrow_d \chi^2(l),$$

independently on θ_0. Hence confidence areas are possible.

Numerical applications : Inference for Generalized Pareto family

The Generalized Pareto Distributions (GPD) $\mu = 0$, a scale parameter σ and a shape parameter ν. They can be defined through their density :

$$f_{\sigma,\nu}(x) = \begin{cases} \frac{1}{\sigma}\left(1+\nu\frac{x}{\sigma}\right)^{-1-1/\nu} 1_{x>0} & \text{if } \nu > 0 \\ \frac{1}{\sigma}\exp\left(\frac{x}{\sigma}\right) 1_{x>0} & \text{if } \nu = 0 \\ \frac{1}{\sigma}\left(1+\nu\frac{x}{\sigma}\right)^{-1-1/\nu} 1_{-\sigma/\nu>x>0} & \text{if } \nu < 0 \end{cases}$$

Samples with size $n = 100$.

$$\begin{cases} \lambda_2 = \frac{\sigma}{(1+\nu)(2+\nu)} \\ \lambda_3 = \lambda_2 \frac{1-\nu}{3+\nu} \\ \lambda_4 = \lambda_2 \frac{(1-\nu)(2-\nu)}{(3+\nu)(4+\nu)} \end{cases}$$

for any $\sigma > 0, \nu \in \mathbb{R}$. These distributions share their first L-moments with those of a GPD with scale and shape parameter σ and ν. This estimation can be compared with classical parametric estimators.

The variance and the skewness of the GPD are given by

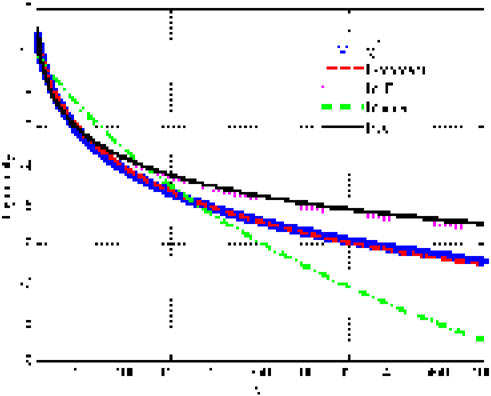
$$\begin{cases} var &= \mathbb{E}[(X-\mathbb{E}[X])^2] = \frac{\sigma^2}{(1-\nu)^2(1-2\nu)} \\ t_3 &= \mathbb{E}\Big[\Big(\frac{X-\mathbb{E}[X]}{\mathbb{E}[(X-\mathbb{E}[X])^2]}\Big)^3\Big] = \frac{2(1+\nu)\sqrt{1-2\nu}}{1-3\nu} \end{cases}$$

var and t_3 respectively exist since $\nu < 1/2$ and $\nu < 1/3$.
On the other hand, the four first L-moments are

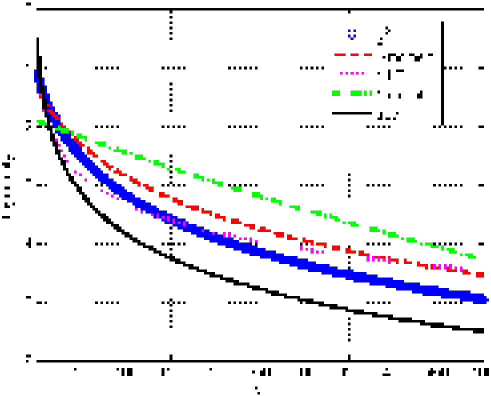
$$\begin{cases} \lambda_1 &= \frac{\sigma}{1-\nu} \\ \lambda_2 &= \frac{\sigma}{(1-\nu)(2-\nu)} \\ \frac{\lambda_3}{\lambda_2} &= \frac{1+\nu}{3-\nu} \\ \frac{\lambda_4}{\lambda_2} &= \frac{(1+\nu)(2+\nu)}{(3-\nu)(4-\nu)} \end{cases}$$

Assuming $\nu < 1$ entails existence of the L-moments.

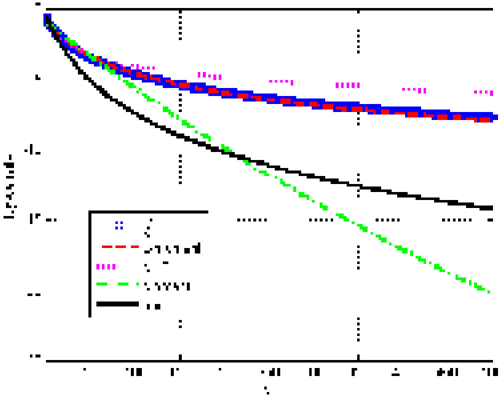
Simulated GPD $\nu = .7$, $\sigma = 3$



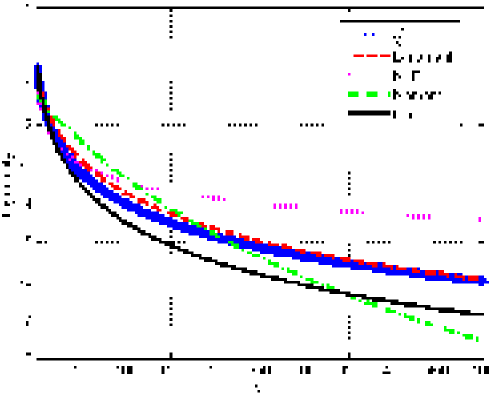
Simulated GPD $\nu = .7, \sigma = 3$, 10% outliers at 300



Simulated GPD $\nu = .1, \sigma = 3$, 10% outliers at 30



Simulated Weibull $\nu = .4$



$\sigma = 3$

LH moments for statistical analysis of extreme events

Linear combinations of high order statistics Wang 1997

$$
\begin{aligned}
\lambda_1^k &:= E\left[X_{k+1,k+1}\right] \\
\lambda_2^k &:= \frac{1}{2}E\left[X_{k+2,k+2} - X_{k+1,k+2}\right] \\
\lambda_3^k &:= \frac{1}{3}E\left[X_{k+3,k+3} - 2X_{k+2,k+3} + X_{k+1,k+1}\right] \\
\lambda_4^k &:= \frac{1}{4}E\left[X_{k+4,k+4} - 3X_{k+3,k+4} + 3X_{k+2,k+4} - X_{k+1,k+4}\right]
\end{aligned}
$$

When $k = 0$ then L-moments.

λ_1^k location of $X_{k+1,k+1}$; λ_2^k =half last gap in a sample of size $k+1$ (spreadness); λ_3^k =asymmetry in the upper tail for large k; λ_4^k =peakedness on the upper part of the distribution.

As k is large characteristics of the upper part of the distribution. Wang (1997): estimators for $\lambda_j^k, 1 \leq j \leq 4$, and LH moments for GEV distributions.Hence similar semiparametric models "close to GEV in the upper tail" for instance.

references

Newey and Smith Higher order properties of GMM..., Econometrica, 2004
Hosking L moments,.., JRSS, 1990
Wang LH moments..., Hydr Research 1997
B Keziou, Inference for models... JSPI, 2012
B Decurninge Inference L-moments , revision IEEE Inf Th