Estimation of the Dependence Structure

Luarens de Haan and Ana Ferriera

November 20, 2020

Chapter 7 of "Extreme value theory: An Introduction"

Presented by Liujun Chen.

Review

The function L is defined by

$$L(x, y) := -\log G_0(\frac{1}{x}, \frac{1}{y}),$$

for $x, y > 0$. And L is connected to the exponent measure v as follows:

$$L(x, y) := v\{(s, t) \in \mathbb{R}_+^2 : s > 1/x \, \text{or} \;\; t > 1/y\}.$$

Review

1. $L(ax, ay) = aL(x, y)$, for all $a, x, y > 0$.
2. $L(x, 0)$=$L(0, x)$=x, for all $x > 0$.
3. $\max(x, y) \leq L(x, y) \leq x + y$, for all $x, y > 0$.
4. If X and Y are independent, then $L(x, y) = x + y$.
5. If $X = Y$ *a.s.*, then $L(x, y) = \max(x, y)$ for $x, y > 0$.
6. L is continous and convex.

Estimation of L

Recall that:

$$\lim_{t\to\infty} t\{1 - F(U_1(\frac{t}{x}), U_2(\frac{t}{y}))\} = L(x, y). \tag{1}$$

Substiute $t = n/k$, (1) can be read as

$$\lim_{n\to\infty} \frac{n}{k}\{1 - F(U_1(\frac{n}{kx}), U_2(\frac{n}{ky}))\} = L(x, y). \tag{2}$$

Relacing F by F_n, $U_1(\frac{n}{kx})$ by $X_{n-[kx]+1,n}$, and $U_2(\frac{n}{ky})$ by $Y_{n-[ky]+1,n}$, we get

$$\hat{L}(x, y) := \frac{1}{k}\sum_{i=1}^{n} 1_{\{X_i \geq X_{n-[kx]+1,n} \text{ or } Y_i \geq Y_{n-[ky]+1,n}\}}. \tag{3}$$

Consistency

Suppose F is in the domian of extreme value distribution G. Let the marginal distributiobn function pf G are exactly $\exp(-(1+\gamma_i x)^{-1/\gamma_i})$ for $i = 1, 2$. Then for $T > 0$ as $n \to \infty, k = k(n) \to \infty, k/n \to 0$,

$$\sup_{0 \leq x, y \leq T} |\hat{L}(x, y) - L(x, y)| \xrightarrow{P} 0.$$

Sketch of the Proof:

1. Prove pointwise convergence.
2. Prove Convergence of the Process.

Asymptotical Normality

Further Assumption:

- Suppose that for some $\alpha > 0$ and for all $x, y > 0$,

$$t\{1 - F(U_1(\frac{t}{x}), U_2(\frac{t}{y}))\} = L(x,y) + O(t^{-\alpha}), \tag{7.2.8}$$

  holds uniformly on the set

$$\{x^2 + y^2 = 1, x \geq 0, y \geq 0\}.$$

- The function L has continous first-order partial derivatives

$$L_1(x,y) := \frac{\partial}{\partial x} L(x,y), \quad \text{and} \quad L_2(x,y) := \frac{\partial}{\partial y} L(x,y).$$

Asymptotical Normality

We first introduce a meansure μ that is closely related to the measure v as follows: for $x, y > 0$,

$$\mu\{(s,t) \in [0,\infty]^2 \setminus \{(\infty,\infty)\} : s < x \text{ or } t < y\}$$
$$:= v\{(s,t) \in [0,\infty]^2 \setminus \{(0,0)\} : s > 1/x \text{ or } t > 1/y\}.$$

Let $D([0,T] \times [0,T])$ be the space of the functions in $[0,T] \times [0,T]$ that are right continous and have finite left-hand limits.

Asymptotical Normality

Then for $k = k(n) \to \infty, k(n) = o(n^{2\alpha/(1+2\alpha)})$, as $n \to \infty$,

$$\sqrt{k}(\hat{L}(x,y) - L(x,y)) \xrightarrow{d} B(x,y),$$

in $D([0, T] \times [0, T])$, for every $T > 0$, where

$$B(x,y) = W(x,y) - L_1(x,y)W(x,0) - L_2(x,y)W(0,y),$$

and W is a continous mean-zero Gaussian process with covariance structure

$$EW(x_1,y_1)W(x_2,y_2) = \mu(R(x_1,y_1) \cap R(x_2,y_2)),$$

with

$$R(x,y) := \{(u,v) \in \mathbb{R}_+^2 : 0 \le u \le x \,\text{or}\, 0 \le v \le y\}.$$

Proposition 7.2.3

Define

$$U_i := 1 - F_1(X_i), \quad \text{and} \quad W_i := 1 - F_2(Y_i),$$

and

$$V_{n,k}(x,y) := \frac{1}{k}\sum_{i=1}^{n} 1_{\{U_i \le kx/n \text{ or } W_i \le ky/n\}}.$$

Then, provided $k \to \infty$, $k/n \to 0$, *as* $n \to \infty$,

$$\sqrt{k}\Big(V_{n,k}(x,y) - \frac{n}{k}\{1 - F(U_1(\frac{n}{kx}), U_2(\frac{n}{ky}))\}\Big) \xrightarrow{d} W(x,y),$$

in $D([0,T] \times [0,T])$, for every $T > 0$.

Proposition 7.2.3

Sketch of the proof:

- Finite-dimensional distributions

    Lyapunov's form of the central limit theorem.

    Cramér Wold theorem

- Tightness

Cramér Wold Theorem

Let $\boldsymbol{X}_n = (X_{n1}, X_{n2}, \ldots, X_{nk})$ and $\boldsymbol{X} = (X_1, X_2, \ldots, X_k)$ be random vectors of dimension k.

Then $\boldsymbol{X}_n$ converges to $\boldsymbol{X}$ if and only if

$$\sum_{i=1}^{k} t_i X_{ni} \xrightarrow{d} \sum_{i=1}^{k} t_i X_i.$$

Tightness

A sequences of random variables is tight if, for all $\varepsilon > 0$, there exists a compact set $K = K(\varepsilon)$ such that

$$\sup_n P(X_n \in K^c) < \varepsilon.$$

Corollary 7.2.4

If Moreove (7.2.8) holds, $k \to \infty, k(n) = o(n^{2\alpha/(1+2\alpha)})$ as $n \to \infty$, then

$$\sqrt{k}\big(V_{n,k}(x,y) - L(x,y)\big) \xrightarrow{d} W(x,y),$$

in $D([0,T] \times [0,T])$, for every $T > 0$.

Sketch of the proof:
Skorohod's Representation.

Skorohod's Representation

Let $\{X_n, n \geq 1\}$ be random variables such that

$$X_n \xrightarrow{d} X \quad \text{as} \quad n \to \infty.$$

Then there exist random variables X' and $\{X'_n, n \geq 1\}$ defined on the Lebesgue probability space, such that

$$X'_n \stackrel{d}{=} X_n \quad \text{for} \quad n \geq 1, \quad X' \stackrel{d}{=} X, \quad and \quad X'_n \stackrel{a.s.}{\to} X' \quad as \quad n \to \infty.$$

Estimation of the Spectral Measure

- In section 7.2, we were concerned with estimating the extremes value distribution G_0 via estimation of the function $L(x, y) := -\log G_0(1/x, 1/y), x, y > 0$.
- In general, $\hat{G}_0 := \exp(-\hat{L}(1/x, 1/y))$ itself is not an extreme value distribution since it is not guaranteed that $\hat{L}$ satisfies the homogeneity property that is valid for the function L:

$$L(ax, ay) = aL(x, y),$$

  for $a, x, y > 0$.
- It is useful to develop an estimatior for G_0 that itself is an extreme value distribution.

Estimation of the Spectral Measure

- This can be done Theorem 6.1.4, which states any finite measure satisfying the side conditions, represented by the distribution function Φ, give rise to an extreme value distribution G_0 via (6.1.31).
- Hence now we focus on the estimation of the spectral measure and in order to do so we have to go back to the origin of this measure.
- We discuss only the spectral measure of Theorem 6.1.14(3) and not the other two, since asymptotic normality has been proved so far only for the third of the spectral measure.

Estimation of the Spectral Measure

Recall that

$$\Phi(\theta) = \mu(E_{1,\theta})$$

with

$$E_{q,\theta} := \{(x, y) \in [0, \infty]^2 \setminus \{(\infty, \infty)\} : x \wedge y < q \,\text{and}\, y/x \leq tan\theta\},$$

for some $q > 0$ and $\theta \in [0, \frac{\pi}{2}]$. Based on the proof of Theorem 6.1.9,

$$\lim_{t \to \infty} tP\Big((1 - F_1(X)) \wedge (1 - F_2(Y)) \leq \frac{1}{t} \quad \text{and} \quad \frac{1 - F_2(Y)}{1 - F_1(X)} \leq tan\theta\Big)$$
$$= \mu(E_{1,\theta}) = \Phi(\theta),$$

for all continuity points θ of Φ.

Estimation of the Dependence Structure

We replace the measure P by its empirical counterpart. We use $R(X_i)$ to denote the rank of the i-th observation $X_i, i = 1, 2, \ldots, n$, among $(X_1, X_2, \ldots, X_n)$.
Taking everything together we get the following estimator for Φ:

$$\hat{\Phi}(\theta) := \frac{1}{k} \sum_{i=1}^{n} 1_{\{R(X_i) \vee R(Y_i) \geq n+1-k \text{ and } n+1-R(Y_i) \leq (n+1-R(X_i)) tan\theta\}}.$$

Estimation of L

Reall that

$$L(x,y) = \int_0^{\pi/2} \{(x(1 \wedge tan\theta)) \vee (y(1 \wedge cot\theta))\}\Phi(d\theta),$$

for $x, y > 0$. Based on the proof of Theorem 7.3.1, the alternative expression for $L(x,y)$ is

$$L(x,y) = x\Phi(\frac{\pi}{2}) + (x \vee y)\int_{\pi/4}^{arctan(y/x)} \Phi(\theta)(\frac{1}{sin^2\theta} \wedge \frac{1}{cos^2\theta})d\theta.$$

This leads to an alternative estimator of the function L with Φ repalced by $\hat{\Phi}$.

Proof of the above page

First, we split the integrate into two parts: the first is from 0 to $\pi/4$, and the second is from $\pi/4$ to $\pi/2$. Without loss of generality, we assume that $x \leq y$, Then we have that
$L(x,y) = \int_0^{\pi/4} \{x \tan\theta \vee y\} \Phi(d\theta) + \int_{\pi/4}^{\pi/2} \{x \vee y \cot\theta\} \Phi(d\theta)$.
Since $x \leq y$, then $\arctan y/x > \pi/4$. Then, we obtain that

$$\begin{aligned} L(x,y) &= \int_0^{\pi/4} y\Phi(d\theta) + \int_{\pi/4}^{\arctan y/x} y\cot\theta\Phi(d\theta) + \int_{\arctan y/x}^{\pi/2} x\Phi(d\theta) \\ &:= I_1 + I_2 + I_3. \end{aligned}$$

It is obvious that $I_1 = y\Phi(\pi/4), I_3 = x\Phi(\pi/2) - x\Phi(\arctan y/x)$. For I_2, we have that

$$I_2 = y\cot\theta\Phi(\theta)\Big|_{\pi/4}^{\arctan y/x} + y\int_{\pi/4}^{\arctan y/x} \frac{1}{\sin^2\theta}\Phi(\theta)d\theta$$

Proof: Continue

It follows that $I_2 = x\Phi(\arctan\theta) - y\Phi(\pi/4)$. Then, we obtain that

$$L(x,y) = x\Phi(\pi/2) + y\int_{\pi/4}^{\arctan y/x} \frac{1}{\sin^2\theta}\Phi(\theta)d\theta.$$

The statement holds for $x \leq y$. Recall that

$$\int_0^{\pi/2}(1\wedge\tan\theta)\Phi(d\theta) = \int_0^{\pi/2}(1\wedge\cot\theta)\Phi(d\theta) = 1.$$

Then the statement also holds for $x \geq y$.

Estimation of L and G

- This estimator is somewhat more complicated than the one in Section 7.2. On the other hand, the present estimator has the advantage that it is homogeneous,

$$\hat{L}_\Phi(ax, ay) = a\hat{L}_\Phi(x, y),$$

  for $a, x, y > 0$.
- Therefore the function

$$\hat{G}_0(x, y) := \exp\big(-\hat{L}_\Phi(1/x, 1/y)\big)$$

  is an estimator of the max-stable distribution G_0.

Consistency: Theorem 7.3.1

Let $k = k(n)$ be a sequence of integers such that $k \to \infty, k/n \to 0, n \to \infty$. Then

$$\hat{\Phi}(\theta) \xrightarrow{p} \Phi(\theta),$$

for $\theta = \pi/2$ and each $\theta \in [0, \pi/2)$ that is a continuity point of Φ. Moreover,

$$\hat{L}_{\Phi}(x, y) \xrightarrow{p} L(x, y)$$

for $x, y \geq 0$.

Corollary 7.3.2

The statement of Theorem 7.3.1 imply the seeming stronger statements

$$\lim_{n\to\infty} P(\lambda(\hat{\Phi}, \Phi) > \epsilon) = 0$$

for each $\epsilon > 0$, where λ is the *Lévy* distance:

$$\lambda(\hat{\Phi}, \Phi) \\ = \inf\{\delta : \hat{\Phi}(\theta - \delta) - \delta \leq \Phi(\theta) \leq \hat{\Phi}(\theta + \delta) + \delta \textit{ for all } 0 \leq \theta \leq \pi/2\}$$

and for all $L > 0$,

$$\sup_{0<\leq x,y \leq L} |\hat{L}_{\Phi}(x, y) - L(x, y)| \xrightarrow{p} 0.$$

A Dependent Coefficient

- Consider a random vector $(X_1, \ldots, X_n)$ with distribution $F \in D(G)$,
- Let $K(t) := K_1(t) + \cdots + K_d(t)$ with $K_i(t) = 1_{\{X_i \geq U_i(t)\}}$,
- Define

$$\begin{aligned}
\kappa :&= \lim_{t\to\infty} E(K(t)|K(t) \geq 1) \\
&= \lim_{t\to\infty} \frac{\sum_{j=1}^{d} P(X_j > U_j(t)}{P\big(\cup_{j=1}^{d} X_j > U_j(t)\big)} \\
&= \frac{L(1,0,\ldots,0) + L(0,1,\ldots,0) + \cdots + L(0,\cdots,0,1)}{L(1,1,\cdots,1)} \\
&= \frac{d}{L(1,1,\cdots,1)} := \frac{d}{L}
\end{aligned}$$

A Dependent Coefficient

- The case of asymptotical independence corresponds to $\kappa = 1$.
- The case of full dependence corresponds to $\kappa = d$,
- Define the following dependence coefficient

$$H := \frac{\kappa - 1}{d - 1} = \frac{d - L}{(d - 1)L},$$

- $H = 0$ is equivalent to asymptotical independence and $H = 1$ to full dependence.
- In $\mathbb{R}^2$ it is somewhat usula to consider the dependence coefficient

$$\begin{aligned} \lambda : &= \lim_{t \to \infty} tP(X_1 > U_1(t), X_2 > U_2(t)) \\ &= 2 - L(1, 1), \end{aligned}$$

- $\lambda = 0$ corresponds to asymptotical independence and $\lambda = 1$ to full dependence in R^2.

A Dependent Coefficient

- However, the extension of λ to higher dimensions does not share this property.
- One example is the random vector (Y_1, Y_1, Y_2) with Y_1, Y_2 *i.i.d.* with common distribution $\exp(-1/x)$.
- The exponent measure is concentrated on the intersection of these sets, that is $\{(x_1, x_2, x_3) \in \mathbb{R}^3 : x_1 = x_2, x_3 = 0\}$ and $\{\{(x_1, x_2, x_3) \in \mathbb{R}^3 : x_1 = x_2 = 0\}$. There is no asymptotical independence.
- But
$$\lim_{t\to\infty} tP(X_1 > U_1(t), X_2 > U_2(t), X_3 > U_3(t)) = 0.$$

A Dependent Coefficient

- Extend Theorem 7.2.2 to the d-dimensional case, we have

$$\sqrt{k}(\hat{L} - L) \xrightarrow{d} W(1) - \sum_{i=1}^{d} L_i(1) W^{(i)}.$$

- Since

$$\hat{H} := \frac{d - \hat{L}(1, 1, \ldots, 1)}{(d-1)\hat{L}(1, 1, \ldots, 1)},$$

by Delta method,

$$\sqrt{k}(\hat{H} - H) \xrightarrow{d} N(0, \frac{d\sigma_L}{(d-1)L^2}).$$

- Howerve, when $H = 0$, the asymptotical variance is zero and hence the result cannot be used to hypothesis test.

Tail Porbability

- Suppose one has independent observations $(X_1, Y_1), (X_2, Y_2), \dots, (X_n, Y_n)$ with distribution function F and suppose that we are interested in estimating the probability

$$1 - F(w, z),$$

where $w > \max_{1 \le i \le n}(X_i)$ and $z > \max_{1 \le i \le n} Y_i$.

- We assume that both marginal distribution of F are $1 - 1/x$, a more general situation will be considered in Chapter 8.
- Assume $F \in D(G)$, $w = w_n \to \infty, z = z_n \to \infty$ and moreover that

$$n(1 - F(w_n, z_n))$$

is bounded.

Tail Porbability

- We further assume for simpicity that $w_n = cr_n$ and $z_n = dr_n$, for some positive sequence $r_n \to \infty$ and c, d positive constants.
- Since $F \in D(G)$

$$p_n^* = 1 - F(w_n, z_n) = 1 - F(cr_n, dr_n) \sim \frac{1}{r_n} L(\frac{1}{c}, \frac{1}{d})$$

.

- A nature estimator is

$$\hat{p}_n^* := \frac{1}{r_n} V_{n,k}(\frac{1}{c}, \frac{1}{d}) = \frac{1}{r_n} \frac{1}{k} \sum_{i=1}^{n} 1_{\{X_i \geq nc/k \text{ or } Y_i \geq nd/k\}}$$

Tail Porbability

- Let us look at the problem how to estimate

$$p_n := P(X > w_n, Y > z_n) = P(X > cr_n, Y > dr_n).$$

- One can try to estimate p_n as before by

$$\frac{1}{r_n}\frac{1}{k}\sum_{i=1}^{n} 1_{\{X_i \geq nc/k \text{ and } Y_i \geq md/k\}}$$

$$= \frac{1}{r_n}\frac{1}{k}\sum_{i=1}^{n} 1_{\{X_i \geq nc/k\}} + \frac{1}{r_n}\frac{1}{k}\sum_{i=1}^{n} 1_{\{Y_i \geq nd/k\}} - \frac{1}{r_n}\frac{1}{k}\sum_{i=1}^{n} 1_{\{X_i \geq nc/k \text{ or } Y_i \geq nd}$$

Tail Porbability

- If we assume that the components of F are $i.i.d.$, the rand-hand side of the above relation , multiplied by r_n , converges to $c^{-1} + d^{-1} - (c^{-1} + d^{-1}) = 0$.
- The problem is that in the case of asymptotic independence we know not only that $P(X > tc \text{ and } Y > td)$ is of lower order than $P(X > tc \text{ or } Y > td)$ as $t \to \infty$, but the theory does not say anything about the asymptotical behaviour of the probability itself.
- So, we need more assumption.

Tail Porbability

- Assume the second-order condition

$$\lim_{t\to\infty} \frac{t(1 - F(tx, ty)) - L(\frac{1}{x}, \frac{1}{y})}{A(t)} = Q(x, y)$$

- In cases of asymptotical independece this second order condition takes a simple form. Taking $x = \infty$ or $y = \infty$ we get

$$\frac{t(1 - F(tx, \infty)) - \frac{1}{x}}{A(t)} \to Q(x, \infty),$$

$$\frac{t(1 - F(\infty, ty)) - \frac{1}{y}}{A(t)} \to Q(\infty, y).$$

Tail Porbability

- These imply

$$\frac{tP(X > tx, Y > ty)}{A(t)} \to P(X > tx) + P(Y > ty) - P(X > tx \text{ or } Y > ty) \\ =: S(x, y). \tag{7.5.7}$$

- $P(X > t \text{ or } Y > t)$ is a regularly varying function of order -1.
- $P(X > t \text{ and } Y > t)$ is a regularly varying function of order $\rho - 1$. In the original papers, the index is writeen as $-1/\eta, \eta \leq 1$. Clearly, if there is no asymptotical independence, $\eta = 1$.
- It is common to write (7.5.7) as

$$\frac{P(X > tx, Y > ty)}{P(X > t, Y > t} = S(x, y).$$

Tail Porbability

- We take

$$\hat{p}_n := (\frac{k}{n} r_n)^{-1/\hat{\eta}} \frac{k}{n} \frac{1}{k} \sum_{i=1}^{n} 1_{\{X_i \geq nc/k, Y_i \geq nd/k\}},$$

  where η is an estimator of η to be discussed later.
- If $\hat{\eta}$ converges to η at a certain rate, then we can prove

$$\frac{\hat{p}_n}{p_n} \xrightarrow{p} 1.$$

Estimation of η

- We now define the residual independence parameter η generally.
- Suppose that for $x, y > 0$,

$$\lim_{t \downarrow 0} \frac{P(1 - F_1(X) < tx, 1 - F_2(Y) < ty)}{P(1 - F_1(X) < t, 1 - F_2(Y) < t)} := S(x, y), \tag{7.6.1}$$

exists and is positive,

- Then $P(1 - F_1(X) < t, 1 - F_2(Y) < t)$ is regularly varing function with index $1/\eta$, for $a, x, y > 0$,

$$S(ax, ay) = a^{1/\eta} S(x, y).$$

Estimation of η

- If there is no symptotical independenc, the index η has to be 1.
- $\eta < 1$ imply asymptotical independence.
- $\eta = 1$ does not imply asymptotical independence.

Estimation of η

- Condition (7.6.1) implies:

$$\lim_{t\downarrow 0} \frac{P(\frac{1}{1-F_1(X)} \wedge \frac{1}{1-F_2(Y)} > tx)}{P(\frac{1}{1-F_1(X)} \wedge \frac{1}{1-F_2(Y)} > t)} = S(\frac{1}{x}, \frac{1}{y}) = x^{-1/\eta} S(1,1) = x^{-1/\eta}.$$

- The probability distribution of the random variables $((1-F_1(X)) \vee (1-F_2(Y))^{-1}$ is regularly with index $-1/\eta$.
- This suggests that we use a Hill-type estimator.

Estimation of η

Define

$$T_i^{(n)} := \frac{1}{\left((1 - F_1^{(n)}(X_i)\right) \vee \left((1 - F_2^{(n)}(Y_i)\right)}.$$

Then Hill-type estimator then becomes

$$\hat{\eta} := \frac{1}{k} \sum_{i=0}^{k-1} \log T_{n-i,n}^{(n)} - \log T_{n-k,n}^{(n)},$$

where $\{T_{j,n}\}$ are the order statistics of $T_i^{(n)}, i = 1, 2, \ldots, n$.

Asymptotical normality

For the proof of Asymptotical normality, we need second order assumption.Assume further:

- $$\lim_{t\downarrow 0} \frac{\frac{P(1-F_1(X)<tx,1-F_2(Y)<ty)}{P(1-F_1(X)<t,1-F_2(Y)<t)} - S(x,y)}{q_1(t)} =: Q(x,y)$$

  exists for all $x, y \geq 0$ with $x + y > 0$.
- We assume that the convergence is uniform on $\{(x,y) \in \mathbb{R}_+^2 : x^2 + y^2 = 1\}$.
- The function S has first-order partial derivates S_x, S_y.
- $\lim_{t\downarrow 0} t^{-1} P(1 - F_(X) < t, 1 - F_2(Y) < t) := l$ exists.

Asymptotical normality

For a sequence $k = k(n)$ of integers with $k \to \infty, k/n \to 0$ and $\sqrt{k} q_1(q^{\leftarrow}(k/n)) \to 0, n \to \infty$,

$$\sqrt{k}(\hat{\eta} - \eta)$$

is asymptotical normal with mean zero and variance

$$\eta^2(1 - l)(1 - 2lS_x(1,1)S_y(1,1).$$