23 augustus 2000

1 Extreme quantile and tail estimation

Suppose F is in the domain of attraction of an extreme–value distribution, i.e. for sequences $a_n > 0$ and b_n we have

$$\lim_{n\to\infty} F^n(a_n x + b_n) = \exp -(1+\gamma x)^{-1/\gamma}$$

with $\gamma \in \mathbb{R}$. Then there exist functions a and b such that $a_n = a(n)$ and $b_n = b(n)$. Moreover we can take $b(t) = U(t) = F^{\leftarrow}(1-1/t)$. Then

$$\lim_{t\to\infty} \frac{U(tx)-U(t)}{a(t)} = \frac{x^\gamma - 1}{\gamma}$$

for $x > 0$. Let k be an intermediate sequence i.e. $k = k(n) \to \infty,\ k(n)/n \to 0$ $(n \to \infty)$.

Proposition 1.1. Suppose the second order condition

$$\lim_{t\to\infty} \frac{\frac{U(tx)-U(t)}{a(t)} - \frac{x^\gamma-1}{\gamma}}{A(t)} = \frac{1}{\rho}\left[\frac{x^{\gamma+\rho}-1}{\gamma+\rho} - \frac{x^\gamma-1}{\gamma}\right]$$

holds for $x > 0$ with $\rho \le 0$ and A a function not changing sign and such that $A(t) \to 0$, $t \to \infty$.

Suppose further $\gamma \neq \rho$, then if $U(\infty) > 0$

$$\lim_{t\to\infty} \frac{\frac{\log U(tx)-\log U(t)}{a(t)/U(t)} - \frac{x^{\gamma_-}-1}{\gamma_-}}{\tilde{A}(t)} = \frac{1}{\rho'}\left[\frac{x^{\gamma_-+\rho'}-1}{\gamma_-+\rho'} - \frac{x^{\gamma_-}-1}{\gamma_-}\right]$$

with $\gamma_- = \min(\gamma, 0),\ \rho' \le 0$ and $\tilde{A}(t) \to 0,\ t \to \infty$.

If $\gamma \neq \rho,\ k = k(n) \to \infty$ and

$$\lim_{n\to\infty} \sqrt{k}\tilde{A}\left(\frac{n}{k(n)}\right) = 0,$$

then for suitable estimators $\hat{\gamma}_{k,n},\ \hat{a}(\frac{n}{k})$ and $\hat{b}(\frac{n}{k})$ the random vector

$$\sqrt{k}\left(\hat{\gamma}_{k,n} - \gamma,\ \frac{\hat{a}(\frac{n}{k})}{a(\frac{n}{k})} - 1,\ \frac{\hat{b}(\frac{n}{k}) - b(\frac{n}{k})}{a(\frac{n}{k})}\right)$$

converges in distribution to a 3–dimensional normal distribution with mean vector zero and known covariance matrix depending on γ (not ρ').

Now we are ready to consider extreme quantile estimation. A $(1-p)$–quantile is called extreme in relation to a sample of size n if $p << 1/n$. In that case, since we want to use asymptotic methods pretending that $n \to \infty$, we are forced to assume that in fact $p = p_n$ and $p_n \to 0,\ n \to \infty$. We want to estimate x_{p_n} with $1 - F(x_{p_n}) = p_n$ or, equivalently, $x_{p_n} = U(1/p_n)$.

Theorem 1.2. *Suppose for some function A with $\lim_{t\to\infty} A(t) = 0, \rho \leq 0$*

$$\lim_{t\to\infty} \frac{\frac{U(tx)-U(t)}{a(t)} - \frac{x^\gamma - 1}{\gamma}}{A(t)} = \frac{1}{\rho}\left[\frac{x^{\gamma+\rho}-1}{\gamma+\rho} - \frac{x^\gamma - 1}{\gamma}\right] \tag{1}$$

(second order condition).
Write $a_n := \frac{k}{np_n}$. If

1. $\rho < 0$
2. $\sqrt{k} A\left(\frac{n}{k}\right) \to 0 \qquad (n \to \infty)$
3. $\log(np_n) = o(\sqrt{k}) \qquad \Big($*or:* $\log p_n = o\left(n^\varepsilon\right), \forall \varepsilon > 0\ \Big)$.

Define

$$\hat{x}_{p_n} := \hat{b}\left(\frac{n}{k}\right) + \hat{a}\left(\frac{n}{k}\right) \cdot \frac{\left(\frac{k}{np_n}\right)^{\hat{\gamma}_{k,n}} - 1}{\hat{\gamma}_{k,n}} \qquad \textit{and } x_{p_n} := U\left(\frac{1}{p_n}\right).$$

Then

$$\sqrt{k(n)} \frac{\hat{x}_{p_n} - x_{p_n}}{a\left(\frac{n}{k}\right) \cdot q_\gamma(a_n)}$$

has asymptotically a normal distribution, mean zero, known variance (depending on γ, not ρ), with

$$q_\gamma(a_n) := \int_1^{a_n} s^{\gamma-1}(\log s) ds.$$

For the proof we need the following lemma:

Lemma 1.3. If (1) holds with $\rho < 0$, then

$$\lim_{\substack{t\to\infty \\ x=x(t)\to\infty}} \frac{\frac{U(tx)-U(t)}{a(t)} \cdot \frac{\gamma}{x^{\gamma}-1} - 1}{A(t)} = \frac{-1}{\rho + \gamma_{-}}$$

Proof.

$$\begin{aligned}
\hat{x}_p - x_p &= \hat{b} + \hat{a} \cdot \frac{a_n^{\hat{\gamma}} - 1}{\hat{\gamma}} - U\left(\frac{1}{p}\right) \\
&= \hat{b} - b + \hat{a} \cdot \frac{a_n^{\hat{\gamma}} - 1}{\hat{\gamma}} - \left(U\left(\frac{1}{p}\right) - U\left(\frac{n}{k}\right)\right) \\
&= \hat{b} - b + \hat{a}\left(\frac{a_n^{\hat{\gamma}} - 1}{\hat{\gamma}} - \frac{a_n^{\gamma} - 1}{\gamma}\right) + (\hat{a} - a) \cdot \frac{a_n^{\gamma} - 1}{\gamma} \\
&- \left\{U\left(\frac{1}{p}\right) - U\left(\frac{n}{k}\right) - a\frac{a_n^{\gamma} - 1}{\gamma}\right\}.
\end{aligned}$$

Hence

$$\sqrt{k}\frac{\hat{x}_p - x_p}{a\left(\frac{n}{k}\right) \cdot q_{\gamma}(a_n)} =$$

$$\begin{array}{lll}
\mathrm{I} & \sqrt{k}\ \dfrac{\hat{b} - b}{a}\dfrac{1}{q_{\gamma}} + & \left(\hat{b}\text{ part}\right) \\
\mathrm{II} & +\dfrac{\hat{a}}{a}\ \left[\dfrac{\sqrt{k}}{q_{\gamma}(a_n)}\left(\dfrac{a_n^{\hat{\gamma}} - 1}{\hat{\gamma}} - \dfrac{a_n^{\gamma} - 1}{\gamma}\right)\right] & (\hat{\gamma}\text{ part}) \\
\mathrm{III} & +\sqrt{k}\ \left(\dfrac{\hat{a}}{a} - 1\right)\dfrac{a_n^{\gamma} - 1}{\gamma q_{\gamma}(a_n)} & (\hat{a}\text{ part}) \\
\mathrm{IV} & +\dfrac{\sqrt{k}}{q_{\gamma}(a_n)}\ \left[\dfrac{U\left(\frac{n}{k} \cdot a_n\right) - U\left(\frac{n}{k}\right)}{a\left(\frac{n}{k}\right)} - \dfrac{a_n^{\gamma} - 1}{\gamma}\right] & \text{(non-random: bias)}
\end{array}$$

It is easy to see that

$$q_{\gamma}(a_n) \sim \begin{cases} \frac{1}{\gamma}a_n^{\gamma}/\log a_n & \gamma > 0 \\ \frac{1}{2}(\log a_n)^2 & \gamma = 0 \\ 1/\gamma^2 & \gamma < 0. \end{cases} \tag{2}$$

Hence I and III are asymptotically normal and we know the joint limit distribution of the limit random variables.

Next consider II which is basically

$$\frac{\sqrt{k}}{q_\gamma(a_n)}\left(\frac{a_n^{\hat\gamma}-1}{\hat\gamma}-\frac{a_n^{\gamma}-1}{\gamma}\right).$$

First note that for alle real x

$$\left|\frac{e^x-1}{x}-1\right|\le e^{|x|}-1. \tag{3}$$

$$\left(\begin{array}{lcl} LHS & = & |\frac{x}{2!}+\frac{x^2}{3!}+\dots\ |\le\frac{|x|}{2!}+\frac{x^2}{3!}+\frac{|x|^3}{4!}+\dots \\ RHS & = & |x|+\frac{x^2}{2!}+\frac{|x|^3}{3!}+\dots \\ & & \text{hence o.k.}\end{array}\right)$$

Next note that

$$\frac{x^\gamma-1}{\gamma}=\int_1^x s^{\gamma-1}ds \quad \text{for all } x>0,\ \gamma\in\mathbb{R}.$$

Hence

$$\frac{a_n^{\hat\gamma}-1}{\hat\gamma}-\frac{a_n^{\gamma}-1}{\gamma}=\int_1^{a_n}\left(s^{\hat\gamma-1}-s^{\gamma-1}\right)ds=\int_1^{a_n}s^{\gamma-1}\left(s^{\hat\gamma-\gamma}-1\right)ds.$$

It follows that

$$\left|\frac{a_n^{\hat\gamma}-1}{\hat\gamma}-\frac{a_n^{\gamma}-1}{\gamma}-(\hat\gamma-\gamma)\int_1^{a_n}s^{\gamma-1}(\log s)ds\right|$$

$$=\left|\int_1^{a_n}s^{\gamma-1}\frac{e^{(\hat\gamma-\gamma)\log s}-1}{(\hat\gamma-\gamma)\log s}(\hat\gamma-\gamma)(\log s)ds-(\hat\gamma-\gamma)\int_1^{a_n}s^{\gamma-1}(\log s)ds\right|$$

$$=\left|(\hat\gamma-\gamma)\int_1^{a_n}s^{\gamma-1}\left[\frac{e^{(\hat\gamma-\gamma)\log s}-1}{(\hat\gamma-\gamma)\log s}-1\right]\log s\ ds\right| \tag{4}$$

$$\le|\hat\gamma-\gamma|\int_1^{a_n}\left(e^{|\hat\gamma-\gamma|\log s}-1\right)\log\ \ ds$$

$$\le|\hat\gamma-\gamma|\left(e^{|\hat\gamma-\gamma|\log a_n}-1\right)\int_1^{a_n}s^{\gamma-1}\log s\ ds.$$

Hence

$$\left| \frac{\sqrt{k}}{q_\gamma(a_n)} \left(\frac{a_n^{\hat\gamma} - 1}{\hat\gamma} - \frac{a_n^{\gamma} - 1}{\gamma} \right) - \sqrt{k}(\hat\gamma - \gamma) \right| \leq$$

$$\leq \sqrt{k}|\hat\gamma - \gamma| \left[\exp\left\{ \sqrt{k}|\hat\gamma - \gamma| \cdot \frac{\log a_n}{\sqrt{k}} \right\} - 1 \right]$$

Conclusion: the limit distribution of part II is the same as that of $\sqrt{k}(\hat\gamma - \gamma)$. Finally part IV:

$$\frac{\sqrt{k}}{q_\gamma(a_n)} \left[\frac{U(a_n \frac{n}{k}) - U(\frac{n}{k})}{a(\frac{n}{k})} - \frac{a_n^\gamma - 1}{\gamma} \right] =$$

$$= \sqrt{k} A(\frac{n}{k}) \cdot \frac{a_n^\gamma - 1}{\gamma q_\gamma(a_n)} \cdot \left[\left\{ \frac{U(a_n \frac{n}{k}) - U(\frac{n}{k})}{a(\frac{n}{k})} \cdot \frac{\gamma}{a_n^\gamma - 1} - 1 \right\} \big/ A(\frac{n}{k}) \right].$$

To finish the proof note (2) and

$$\frac{a_n^\gamma - 1}{\gamma} \sim \begin{cases} \frac{1}{\gamma} a_n^\gamma & \gamma > 0 \\ \log a_n & \gamma = 0 \\ -1/\gamma & \gamma < 0, \end{cases} \tag{5}$$

hence $(a_n^\gamma - 1)/\{\gamma q_\gamma(a_n)\}$ is bounded for $\gamma \in \mathbb{R}$. So by assumption 2 and Lemma 1.3 part IV converges to zero. □

Comment By checking the components of the asymptotic variance in the Theorem, one sees that the uncertainty in the estimation of x_p is for a large part caused by the uncertainty in the estimation of γ and not so much in the estimation of $b(\frac{n}{k})$ or $a(\frac{n}{k})$.

Special case $\gamma > 0$:

Theorem 1.4. *Suppose for some function A with $\lim_{t\to\infty} A(t) = 0$ and $\rho \leq 0$*

$$\lim_{t\to\infty} \frac{\frac{U(tx)}{U(t)} - x^\gamma}{A(t)} = x^\gamma \cdot \frac{x^\rho - 1}{\rho}.$$

Write $a_n := \frac{k}{np_n}$. If

1. $\rho < 0$

2. $\sqrt{k}A(\frac{n}{k}) \to 0 \qquad (n \to \infty)$

3. $\log np_n = o(\sqrt{k}) \qquad$ *(or* $\log p_n = o(n^\varepsilon), \forall \varepsilon > 0$*).*

Define

$$\hat{x}_{p_n} := X_{n-k,n} \cdot (\frac{k}{np_n})^{\hat{\gamma}_{k,n}} \qquad \textit{and} \quad x_{p_n} := U(\frac{1}{p_n}).$$

Then

$$\frac{\sqrt{k}}{\log a_n}\left(\frac{\hat{x}_p}{x_p} - 1\right)$$

has asymptotically a normal distribution, mean zero, known variance (depending on γ*, not* ρ*).*

Corollary 1.5. Hence in this case

$$\frac{\hat{x}_p}{x_p} \xrightarrow{p} 1 \qquad (n \to \infty).$$

Dual problem: given a large value x, estimate $1 - F(x)$. As explained before we choose

$$1 - \hat{F}_n(x_n) := \frac{k}{n}\max\left\{0, \left[1 + \hat{\gamma}_n \frac{x_n - \hat{b}(\frac{n}{k})}{\hat{a}(\frac{n}{k})}\right]^{-1/\hat{\gamma}_n}\right\}. \tag{6}$$

Assume that there is a sequence $c_n > 0$ with

$$x_n = b(\frac{n}{k}) + a(\frac{n}{k}) \cdot \frac{c_n^\gamma - 1}{\gamma}. \tag{7}$$

Typically $c_n \to \infty$ when $n \to \infty$.

Theorem 1.6. *Assume the second order condition. If* $\rho < 0$*,*

1. $k^{-1/2}(c_n^{-\gamma_-} - 1)/(-\gamma_-) \to 0$ *and*

2. $\sqrt{k}A(\frac{n}{k})(\log c_n) \to 0 \qquad (n \to \infty),$

then

$$\frac{\sqrt{k}}{q_{\hat{\gamma}_n}(\hat{c}_n)\left\{\frac{n}{k}\left(1 - \hat{F}_n(x_n)\right)\right\}^{\hat{\gamma}_n}}\left(\frac{1 - \hat{F}_n(x_n)}{1 - F(x_n)} - 1\right)$$

has asymptotically a normal distribution, mean zero, known variance (depending on γ*, not* ρ*), with* q_γ *as before and*

$$\hat{c}_n := \frac{k}{n}\left\{1 - \hat{F}_n(y_n)\right\}^{-1}.$$

For the proof we need the following Lemma:

Lemma 1.7. *If the second order condition holds, then*

$$\lim_{\substack{n\to\infty \\ c_n\to\infty}} \frac{\frac{n}{k}\left\{1-F\left(b(\frac{n}{k})+a(\frac{n}{k})\frac{c_n^\gamma-1}{\gamma}\right)\right\}\left(1+\gamma\frac{c_n^\gamma-1}{\gamma}\right)^{1/\gamma}-1}{A(\frac{n}{k})} = \frac{-1}{\rho+\gamma_-}.$$

Remark 1.8. The limit random variable here is minus the one in the dual case as could be expected from the Bahadur-Kiefer representation.

For the proof we need some Lemmas: throughout we assume the stated conditions.

Lemma 1.9. *Write* $\hat{x}_n := \hat{b}(\frac{n}{k}) + \hat{a}(\frac{n}{k})(c_n^{\hat{\gamma}}-1)/\hat{\gamma}$.

$$W_n := \frac{\sqrt{k}(\hat{x}_n - x_n)}{a(\frac{n}{k})q_\gamma(c_n)} \xrightarrow{d} W, \text{ normally distributed.}$$

Proof.

$$\frac{\hat{x}_n - x_n}{a(\frac{n}{k})} = \frac{\hat{b}(\frac{n}{k})-b(\frac{n}{k})}{a(\frac{n}{k})} + \frac{\hat{a}(\frac{n}{k})}{a(\frac{n}{k})}\left(\frac{c_n^{\hat{\gamma}}-1}{\hat{\gamma}} - \frac{c_n^\gamma-1}{\gamma}\right)$$

$$+\left(\frac{\hat{a}(\frac{n}{k})}{a(\frac{n}{k})}-1\right)\cdot\frac{c_n^\gamma-1}{\gamma}.$$

The rest is as before. □

Lemma 1.10.

$$\frac{nc_n^{\gamma+1}}{q_\gamma(c_n)\sqrt{k}}\left(1-\hat{F}_n(x_n)-\frac{k}{nc_n}\right) \xrightarrow{d} W.$$

Proof.

$$\frac{nc_n}{k}\left(1-\hat{F}(x_n)\right) = c_n \max\left(0, \left\{1+\hat{\gamma}\frac{x_n-\hat{b}}{\hat{a}}\right\}^{-1/\hat{\gamma}}\right)$$

$$= c_n \max\left(0, \left\{1+\hat{\gamma}\left[\frac{x_n-\hat{x}_n}{\hat{a}}+\frac{c_n^{\hat{\gamma}}-1}{\hat{\gamma}}\right]\right\}^{-1/\hat{\gamma}}\right)$$

$$= c_n \max\left(0, \left\{c_n^{\hat{\gamma}} + \hat{\gamma}\frac{x_n - \hat{x}_n}{\hat{a}}\right\}^{-1/\hat{\gamma}}\right) = \max\left(0, \left\{1 + \hat{\gamma}c_n^{-\hat{\gamma}} \cdot \frac{x_n - \hat{x}_n}{\hat{a}}\right\}^{-1/\hat{\gamma}}\right)$$

$$= \max\left(0, \left\{1 + \hat{\gamma}\frac{q_\gamma(c_n)}{c_n^{\hat{\gamma}}\sqrt{k}}W_n\right\}^{-1/\hat{\gamma}}\right).$$

The assumptions imply $c_n^{\hat{\gamma}} \sim c_n^{\gamma}$ and $\frac{q_\gamma(c_n)}{c_n^{\gamma}\sqrt{k}} \to 0$ $(n \to \infty)$.
Hence

$$\frac{nc_n}{k}\left(1 - \hat{F}(x_n)\right) - 1 \sim \frac{q_\gamma(c_n)}{c_n^{\hat{\gamma}}\sqrt{k}}W_n.$$

□

Corollary 1.11.

$$\frac{nc_n}{k}\left(1 - \hat{F}(x_n)\right) \xrightarrow{p} 1.$$

Lemma 1.12.

$$\lim_{n\to\infty} \frac{\sqrt{k}\cdot c_n^{\gamma}}{q_\gamma(c_n)}\left[\frac{nc_n}{k}\left\{1 - F(x_n)\right\} - 1\right] = 0$$

Proof.

$$\frac{\frac{nc_n}{k}\left\{1 - F(x_n)\right\} - 1}{A(\frac{n}{k})} =$$

$$\frac{\frac{n}{k}\left\{1 - F\left(b(\frac{n}{k}) + a(\frac{n}{k})\frac{c_n^{\gamma}-1}{\gamma}\right)\right\}\left(1 + \gamma\frac{c_n^{\gamma}-1}{\gamma}\right)^{1/\gamma} - 1}{A(\frac{n}{k})} \longrightarrow \frac{-1}{\rho + \gamma_-}$$

by Lemma 1.7.
Next note that by assumption $\frac{\sqrt{k}\cdot c_n^{\gamma}}{q_\gamma(c_n)}A(\frac{n}{k}) \to 0 \qquad (n \to \infty)$. □

Proof of the Theorem: combination of the results of Lemma 1.9, 1.10, and 1.12.