Statistica Sinica: Supplement **24** (201?), 000-000

Distinguishing Forms of Asymptotic Dependence in Heavy Tailed Data

Tiandong Wang

Fudan University, Shanghai Academy of Artificial Intelligence for Science

Sidney I. Resnick

Cornell University

Supplementary Material

This supplement contains additional simulation results and proof details for theorems in the main document.

S1 Additional simulation results

In this section, we provide more simulation experiments to examine the finite sample behavior of the proposed test.

Example 1. First, consider the same simulation setup as in Section 5.2. Set $a = 0.25$, $b = 0.75$. Suppose $R_1 \sim \text{Pareto}(2)$, $R_2 \sim \text{Pareto}(6)$, $Z \sim \text{Beta}(0.1, 0.1)$, $\Theta_2 \sim \text{Unif}([0, 1] \setminus [a, b])$, and $B \sim \text{Bernoulli}(0.5)$. Assume the random variables are all independent, and let $\Theta_1 = a + (b - a)Z^2$. Suppose also that

$$X = BR_1\Theta_1 + (1-B)R_2\Theta_2$$

$$Y = BR_1(1-\Theta_1) + (1-B)R_2(1-\Theta_2).$$

Then we generate $n = 1{,}000$ iid observations from the distribution of (X, Y), and the minimum distance method chooses $k(n) = 127 \approx \lceil 8.5n^{0.39}\rceil$, which is further confirmed by the stable shape of the Hill plot (cf. the left panel of Figure 1). Histogram of the angles is in the middle panel of Figure 1 and illustrates the dependence structure. To estimate $[a, b]$, we again solve the optimization problem in Eq. (5.27) with $\lambda = 1$. Also, we generate 100 simulated trials, each of which consists of $n = 1{,}000$ iid observations, to assess the consistency of the estimated $\hat{a}$ and $\hat{b}$, giving small MSE values of 1.04×10^{-3} and 2.24×10^{-8}, respectively.

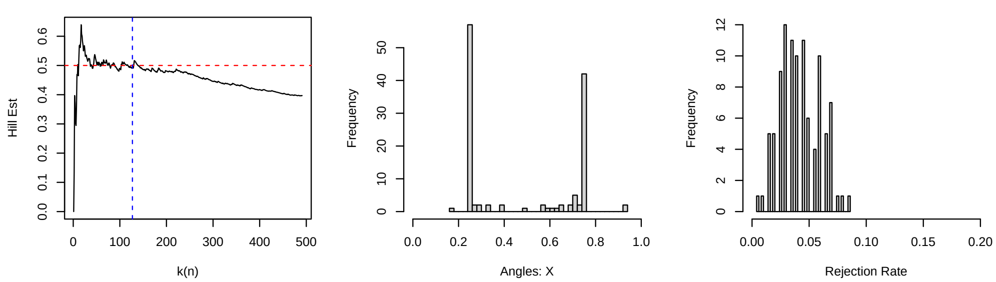

Figure 1: Left: Hill plot of $|x| + |y|$, and the red line denotes $1/\alpha = 0.5$. Middle: Histogram of Θ with $n = 1{,}000$ and $k(n) = 127$. Right: Histogram of rejection rates of $H_0^{(1)} : S([\hat{a}, \hat{b}]) = 1$ by $\widehat{D}_m$ among 100 simulated trials.

Next, we set $m = \lceil 5n/k(n)\rceil$, $k(m) = \lceil 2m^{0.39}\rceil$, and for each of the 100 simulated samples, generate $B = 200$ bootstrap resamples to individually test $H_0^{(i)}$, $i = 1, 2, 3$. Among the 100 trials, $H_0^{(1)}$ is rejected 29% of the time, and $H_0^{(2)}$ is rejected 49% of the time. The rejection rate of $H_0^{(1)}$ for each single trial is plotted in the right panel of Figure 1, which reveals a similar shape as in the right panel of Figure 3, and the average rejection rate is 0.0423 with a standard deviation of 0.0175. In addition, $H_0^{(3)}$ is rejected 19% of the time, suggesting an increase in the type I error due to the decrease in n.

Example 2. Now we consider the power analysis when $n = 1{,}000$, and suppose we have the same setup as in Section 5.2.2. We again generate 100 simulated samples with $n = 1{,}000$, and for $k(n) = 145$ (suggested by both the minimum distance method and the Hill plot in the left panel of Figure 2), the middle panel of Figure 2 gives the histogram of angles from one specific sample. The boxplot in the right panel of Figure 2 gives distributions of the estimated $\hat{a}$ and $\hat{b}$ using $\lambda = 1$. Similar to the right panel of Figure 4, under asymptotic weak dependence, estimated values of $\hat{b}$ vary a lot and are different from the true value of 1.

Next, for each sample, we obtain $B = 200$ resamples of size $m = \lceil 5n/k(n) \rceil$, and set $k(m) = \lceil 2m^{0.39} \rceil$. The rejection rates for $H_0^{(i)}$, $i = 1, 2, 3$, are 27%, 61%, and 23%, respectively. We see that the power of the asymptotic full dependence ($H_0^{(2)}$) test drops due to the decreased sample size, whereas the power of testing $H_0^{(1)}$ and $H_0^{(3)}$ remains the same.

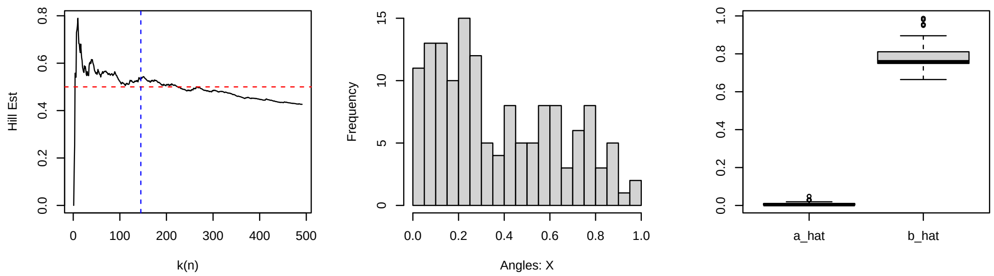

Figure 2: Left: Hill plot of $|x|+|y|$, and the red line denotes $1/\alpha = 0.5$. Middle: Histogram of Θ_1 with $n = 1{,}000$ and $k(n) = 145$. Right: Boxplot of $\hat{a}$ and $\hat{b}$ among 100 simulated trials.

S2 A functional central limit theorem

We rely on the functional central limit theorem given in (Pollard, 1990, Theorem 10.6).

Theorem S2.1. Consider the triangular array $\{f_{n,i}(t) : t \in T\}$ with envelop function $F_{n,i}$, independent within each row. Suppose also that $\{f_{n,i}\}$ satisfy

(i) $\{f_{n,i}\}$ are manageable;

(ii) For $X_n(t) = \sum_i (f_{n,i}(t) - \mathbb{E}(f_{n,i}(t)))$, $H(s,t) = \lim_{n\to\infty} \mathbb{E}[X_n(t)X_n(s)]$ exists for every $s, t \in T$;

(iii) The envelope function satisfies $\limsup_{n\to\infty} \mathbb{E}(F_{n,i}^2) < \infty$, and

$$\sum_i \mathbb{E}\left(F_{n,i}^2 \mathbf{1}_{\{F_{n,i}>\eta\}}\right) \to 0,$$

for each $\eta > 0$;

(iv) Let $\rho_n(s,t) = \left(\sum_i \mathbb{E}\left(f_{n,i}(t) - f_{n,i}(s)\right)^2\right)^{1/2}$, then the limit $\rho(s,t) = \lim_{n\to\infty} \rho_n(s,t)$ is well-defined, and for deterministic sequences $\{s_n\}$, $\{t_n\}$, if $\rho(s_n,t_n) \to 0$, then $\rho_n(s_n,t_n) \to 0$.

Then X_n converges to a Gaussian process with zero mean and covariance given by H.

S3 Proof of Theorem 2.1

For $r > 0$, $I \subset [0,1]$ such that $\chi(\partial((r,\infty)\times I)) = 0$,

$$\chi((r,\infty)\times I) = \lim_{t\to\infty} \frac{t\mathbb{P}[R/b(t) > r, \Theta \in I] - r^{-\alpha}S(I)}{A(t)}$$

and if $S(I) = 0$, this is

$$= \lim_{t\to\infty} \frac{t\mathbb{P}[R/b(t) > r, \Theta \in I]}{A(t)}$$

Set $U(t) := t/A(t)$, $b_0 := b \circ U^{\leftarrow}$ and after a change of variable, the proof of Eq. (2.14) is complete.

S4 Proof of Theorem 3.1

We start by showing that for intermediate sequence $\{k(n)\}$ satisfying Eq. (3.18),

$$W_n(y) := \sqrt{k(n)}\left(\frac{1}{k(n)}\sum_{i=1}^n \left(1 + \frac{d\big(\boldsymbol{Z}_i, \mathbb{C}_{a,b}\big)}{R_{(k(n))}}\right)\mathbf{1}_{\left\{\frac{R_i}{b(n/k(n))} > y\right\}} - y^{-\alpha}\right)$$
$$\Rightarrow W(y^{-\alpha}), \tag{S4.1}$$

in $D(0,\infty)$, where $W(\cdot)$ is a standard Brownian motion.

We begin by showing that the sequence of processes in the variable y satisfy

$$\frac{1}{\sqrt{k(n)}}\sum_{i=1}^{n}\frac{d\big(\boldsymbol{Z}_i,\mathbb{C}_{a,b}\big)}{b(n/k(n))}\mathbf{1}_{\left\{\frac{R_i}{b(n/k(n))}>y\right\}}\Rightarrow 0,\qquad \text{in } D(0,\infty), \tag{S4.2}$$

and it is here that HRV and assumption $[a,b]\subsetneq[0,1]$ is used. We have that

$$\begin{aligned}
&\frac{1}{k(n)}\sum_{i=1}^{n}\frac{d\big(\boldsymbol{Z}_i,\mathbb{C}_{a,b}\big)}{b(n/k(n))}\mathbf{1}_{\left\{\frac{R_i}{b(n/k(n))}>y\right\}}\\
&=\frac{b_0(n/k(n))}{b(n/k(n))}\frac{1}{k(n)}\sum_{i=1}^{n}\frac{d\big(\boldsymbol{Z}_i,\mathbb{C}_{a,b}\big)}{b_0(n/k(n))}\mathbf{1}_{\left\{\frac{R_i}{b(n/k(n))}>y\right\}}\\
&\le\frac{b_0(n/k(n))}{b(n/k(n))}\frac{1}{k(n)}\sum_{i=1}^{n}\frac{d\big(\boldsymbol{Z}_i,\mathbb{C}_{a,b}\big)}{b_0(n/k(n))}\Bigg(\mathbf{1}_{\left\{\frac{R_i}{b(n/k(n))}>y,\,d\left(\boldsymbol{Z}_i,\mathbb{C}_{a,b}\right)>b_0(n/k(n))\epsilon\right\}}\\
&\qquad+\mathbf{1}_{\left\{\frac{R_i}{b(n/k(n))}>y,\,d\left(\boldsymbol{Z}_i,\mathbb{C}_{a,b}\right)\le b_0(n/k(n))\epsilon\right\}}\Bigg)\\
&\le\frac{b_0(n/k(n))}{b(n/k(n))}\left(\frac{1}{k(n)}\sum_{i=1}^{n}\frac{d\big(\boldsymbol{Z}_i,\mathbb{C}_{a,b}\big)}{b_0(n/k(n))}\mathbf{1}_{\left\{\frac{d\left(\boldsymbol{Z}_i,\mathbb{C}_{a,b}\right)}{b_0(n/k(n))}>\epsilon\right\}}+\frac{\epsilon}{k(n)}\sum_{i=1}^{n}\mathbf{1}_{\left\{\frac{R_i}{b(n/k(n))}>y\right\}}\right)\\
&=A+B.
\end{aligned}$$

To handle B observe for each fixed $y>0$, that the monotone function in y,

$$\frac{1}{k(n)}\sum_{i=1}^{n}\mathbf{1}_{\left\{\frac{R_i}{b(n/k(n))}>y\right\}}\Rightarrow y^{-\alpha}$$

using, for example (Resnick, 2007, Theorem 5.3(ii), p. 139). Therefore, when $k(n)$ satisfies Eq. (3.18), we have $\sqrt{k(n)}B\Rightarrow 0$ in $D(0,\infty)$.

For A we claim

$$\frac{1}{k(n)}\sum_{i=1}^{n}\frac{d\big(\boldsymbol{Z}_i,\mathbb{C}_{a,b}\big)}{b_0(n/k(n))}\mathbf{1}_{\left\{\frac{d\left(\boldsymbol{Z}_i,\mathbb{C}_{a,b}\right)}{b_0(n/k(n))}>\epsilon\right\}}=O_p(1),$$

since for any $M > 0$

$$\mathbb{P}\left[\frac{1}{k(n)}\sum_{i=1}^{n}\frac{d\big(\boldsymbol{Z}_i,\mathbb{C}_{a,b}\big)}{b_0(n/k(n))}\mathbf{1}_{\left\{\frac{d\left(\boldsymbol{Z}_i,\mathbb{C}_{a,b}\right)}{b_0(n/k(n))}>\epsilon\right\}}>M\right]$$

$$\leq\frac{1}{M}\frac{n}{k(n)}\mathbb{E}\left(\frac{d\big((X_1,Y_1),\mathbb{C}_{a,b}\big)}{b_0(n/k(n))}\mathbf{1}_{\left\{\frac{d\left((X_1,Y_1),\mathbb{C}_{a,b}\right)}{b_0(n/k(n))}>\epsilon\right\}}\right)$$

and because $\alpha_0 > 1$ we may apply Karamata's theorem on integration to get convergence, as $n \to \infty$ to

$$\to\frac{1}{M}\int_{\epsilon}^{\infty}x\nu_{\alpha_0}(dx)<\infty.$$

Therefore $\sqrt{k(n)}A \Rightarrow 0$ in $D(0,\infty)$. This proves (S4.1).

Since $R_{(k(n))}/b(n/k(n)) \xrightarrow{p} 1$ (eg. (Resnick, 2007, p. 82)), we also have

$$\frac{1}{\sqrt{k(n)}}\sum_{i=1}^{n}\frac{d\big(\boldsymbol{Z}_i,\mathbb{C}_{a,b}\big)}{R_{(k(n))}}\mathbf{1}_{\left\{\frac{R_i}{b(n/k(n))}>y\right\}}\Rightarrow 0,\qquad \text{in } D(0,\infty).$$

So to prove (S4.1) it remains to verify

$$\sqrt{k(n)}\left(\frac{1}{k(n)}\sum_{i=1}^{n}\mathbf{1}_{\left\{\frac{R_i}{b(n/k(n))}>y\right\}}-y^{-\alpha}\right)\Rightarrow W(y^{-\alpha}),\tag{S4.3}$$

in $D(0,\infty)$. The regular variation of $P[R_1 > x]$ implies ((Resnick, 2007, Theorem 9.1, p. 292) or de Haan and Ferreira (2006)) that

$$\sqrt{k(n)}\left(\frac{1}{k(n)}\sum_{i=1}^{n}\mathbf{1}_{\left\{\frac{R_i}{b(n/k(n))}>y\right\}}-\frac{n}{k(n)}\mathbb{P}[R_1/b(n/k)>y]\right)\Rightarrow W(y^{-\alpha}),$$

in $D(0,\infty)$ and the 2RV assumption in Eq. (2.8) marginalized to the distribution of R_1 and the choice of $k(n)$ in Eq. (3.18) imply

$$\sqrt{k(n)}\left(\frac{n}{k(n)}\mathbb{P}\left(\frac{R_1}{b(n/k(n))}>y\right)-y^{-\alpha}\right)\to 0,$$

locally uniformly in y. This gives (S4.3) and completes the proof of (S4.1).

Apply the composition map $(x(t),c)\mapsto x(ct)$ from $D(0,\infty)\times(0,\infty)\mapsto D(0,\infty)$ to (S4.1) in the form

$$\left(W_n(y),\frac{R_{(k(n))}}{b(n/k)}\right)\mapsto W_n\left(\frac{R_{(k(n))}}{b(n/k)}y\right)\Rightarrow W(y^{-\alpha})$$

to get

$$\sqrt{k(n)}\left(\frac{1}{k(n)}\sum_{i=1}^{n}\left(1+\frac{d\big(\boldsymbol{Z}_i,\mathbb{C}_{a,b}\big)}{R_{(k(n))}}\right)\mathbf{1}_{\left\{\frac{R_i}{R_{(k(n))}}>y\right\}}-\big(\frac{R_{(k(n))}}{b(n/k(n))}y\big)^{-\alpha}\right)$$
$$\Rightarrow W(y^{-\alpha}). \tag{S4.4}$$

Couple this with a Vervaat inversion of (S4.3) ((de Haan and Ferreira, 2006, p. 357) or (Resnick, 2007, p. 57)). The inversion yields in $D(0,\infty)$

$$\sqrt{k(n)}\left(\big(\frac{R_{([k(n)t])}}{b(n/k(n))}\big)^{-\alpha}-t\right)\Rightarrow -W(t),\quad \text{in } D(0,\infty), \tag{S4.5}$$

and the convergence is joint with the one in (S4.4). Combining (S4.4) with (S4.5) gives

$$\sqrt{k(n)}\Big(\frac{1}{k(n)}\sum_{i=1}^{n}\left(1+\frac{d\big(\boldsymbol{Z}_i,\mathbb{C}_{a,b}\big)}{R_{(k(n))}}\right)\mathbf{1}_{\left\{\frac{R_i}{R_{(k(n))}}>y\right\}}-y^{-\alpha}\Big)$$
$$\Rightarrow W(y^{-\alpha})-y^{-\alpha}W(1). \tag{S4.6}$$

Finally apply the mapping

$$x\mapsto\int_1^\infty\frac{x(s)}{s}\mathrm{d}s$$

to (S4.4) using justifications similar to (Resnick, 2007, Section 9.1). This yields the asymptotic normality of D_n under the null hypothesis in $H_0^{(1)}$,

$$\sqrt{k(n)}\left(\frac{1}{k(n)}\sum_{i=1}^{n}\left(1+\frac{d\big(\boldsymbol{Z}_i^*,\mathbb{C}_{a,b}\big)}{R_{(k(n))}}\right)\log\frac{R_{(i)}}{R_{(k(n))}}-\frac{1}{\alpha}\right)$$
$$\Rightarrow\frac{1}{\alpha}\left(\int_0^1\frac{W(s)}{s}\mathrm{d}s-W(1)\right)\stackrel{d}{=}\frac{1}{\alpha}W(1)\sim N(0,1/\alpha^2).$$

S5 Proof of Theorem 4.1

Proceeds in steps.

(1) To prove the convergence in Eq. (4.23) under $H_0^{(2)}$, we first show that in $D(0,\infty)$,

$$\frac{\sqrt{k(n)}}{\theta_0}\left(\frac{1}{k(n)}\sum_{i=1}^{n}\Theta_i\epsilon_{R_i/b(n/k(n))}(y,\infty)-\theta_0 y^{-\alpha}\right)\Rightarrow W(y^{-\alpha}). \qquad \text{(S5.7)}$$

Decompose the left side of (S5.7) as

$$\frac{1}{\theta_0}\frac{1}{\sqrt{k(n)}}\sum_{i=1}^{n}(\Theta_i-\theta_0)\,\epsilon_{R_i/b(n/k(n))}(y,\infty)$$
$$+\sqrt{k(n)}\left(\frac{1}{k(n)}\sum_{i=1}^{n}\epsilon_{R_i/b(n/k(n))}(y,\infty)-y^{-\alpha}\right)=:B1+B2. \qquad \text{(S5.8)}$$

From full dependence, we have $\mathbb{C}_{a,b}\equiv\{(x,y)\in\mathbb{R}_+^2:y=(1/\theta_0-1)x\}$. Remember $\Theta_i=X_i/R_i=X_i/(X_i+Y_i)$ and then $|B1|$ is bounded by

$$\begin{aligned}&\frac{1}{\theta_0}\frac{1}{\sqrt{k(n)}}\sum_{i=1}^{n}\left|\frac{X_i}{R_i}-\theta_0\right|\epsilon_{R_i/b(n/k(n))}(y,\infty)\\ &=\frac{1}{\theta_0}\frac{1}{\sqrt{k(n)}}\sum_{i=1}^{n}\theta_0\frac{|Y_i-X_i(\theta_0^{-1}-1)|}{R_i}\mathbf{1}_{\{R_i>b(n/k(n))y\}}\\ &=\frac{1}{\sqrt{k(n)}}\sum_{i=1}^{n}\frac{d(\boldsymbol{Z}_i,\mathbb{C}_{a,b})}{b(n/k(n))y}\mathbf{1}_{\{R_i>b(n/k(n))y\}}\Rightarrow 0\end{aligned} \qquad \text{(S5.9)}$$

in $D(0,\infty)$, since (S4.2) is still applicable. As in (S4.3), the second term $B2$ in (S5.8) converges weakly in $D(0,\infty)$ to $W(y^{-\alpha})$ under the 2RV condition for $\mathbb{P}[R_1 > x]$, thus completing the proof of (S5.7).

Combine (S5.7) with

$$\frac{R_{(k(n))}}{b(n/k(n))} \xrightarrow{p} 1,$$

to get joint convergence in $D(0,\infty) \times \mathbb{R}_+$. Applying the composition map $(x(t), c) \mapsto x(ct)$ from $D(0,\infty) \times (0,\infty) \mapsto D(0,\infty)$, we get in $D(0,\infty)$,

$$\frac{\sqrt{k(n)}}{\theta_0}\left(\frac{1}{k(n)}\sum_{i=1}^{n}\Theta_i \epsilon_{R_i/R_{(k(n))}}(y,\infty) - \theta_0\left(\frac{R_{(k(n))}}{b(n/k(n))}y\right)^{-\alpha}\right) \Rightarrow W(y^{-\alpha}). \tag{S5.10}$$

Apply Vervaat's inversion again as in (S4.5) and (S4.6), we again conclude

$$\sqrt{k(n)}\left(\frac{1}{k(n)\theta_0}\sum_{i=1}^{n}\Theta_i \epsilon_{R_i/R_{(k(n))}}(y,\infty) - y^{-\alpha}\right) \Rightarrow W(y^{-\alpha}) - y^{-\alpha}W(1). \tag{S5.11}$$

(2) Next, we need to justify application of the map

$$x \mapsto \int_1^\infty \frac{x(s)}{s}\mathrm{d}s \tag{S5.12}$$

to (S5.11), which, if justified, leads to

$$\sqrt{k(n)}\left(\frac{1}{k(n)\theta_0}\sum_{i=1}^{k(n)}\Theta_i^* \log\frac{R_{(i)}}{R_{(k(n))}} - \frac{1}{\alpha}\right) \Rightarrow \int_1^\infty \frac{W(y^{-\alpha})}{y}\mathrm{d}y - \frac{1}{\alpha}W(1) \sim \frac{1}{\alpha}N(0,1). \tag{S5.13}$$

The justification for applying the map (S5.12) is somewhat standard (Proposition 9.1 of Resnick (2007)) and is deferred to Section S5.

(3) From (S5.13),

$$\frac{1}{k(n)}\sum_{i=1}^{k(n)}\Theta_i^*\log\frac{R_{(i)}}{R_{(k(n))}}\xrightarrow{p}\frac{\theta_0}{\alpha}, \tag{S5.14}$$

which suggests comparing

$$\begin{aligned}&\sqrt{k(n)}\Big(T_n-\frac{1}{\theta_0}\frac{1}{k(n)}\sum_{i=1}^{k(n)}\Theta_i^*\log\frac{R_{(i)}}{R_{(k(n))}}\Big) \qquad\qquad (\text{S5.15})\\ &\qquad=\sqrt{k(n)}\Big(\frac{1}{\frac{1}{k(n)}\sum_{i=1}^{k(n)}\Theta_i^*}-\frac{1}{\theta_0}\Big)\frac{1}{k(n)}\sum_{i=1}^{k(n)}\Theta_i^*\log\frac{R_{(i)}}{R_{(k(n))}}\end{aligned}$$

and applying (S5.14), this is

$$=\sqrt{k(n)}\Big(\frac{1}{\frac{1}{k(n)}\sum_{i=1}^{k(n)}\Theta_i^*}-\frac{1}{\theta_0}\Big)O_p(1)=\sqrt{k(n)}\left(\frac{\theta_0-\frac{1}{k(n)}\sum_{i=1}^{k(n)}\Theta_i^*}{\frac{1}{k(n)}\sum_{i=1}^{k(n)}\Theta_i^*\theta_0}\right)O_p(1). \tag{S5.16}$$

Now

$$\frac{1}{k(n)}\sum_{i=1}^{n}\epsilon_{(\Theta_i,R_i/R_{(k(n))})}\Rightarrow S\times\nu_\alpha=\epsilon_{\theta_0}\times\nu_\alpha$$

in $\mathbb{M}([0,1]\times\mathbb{R}_+\setminus\{0\})$ (eg. (Resnick, 2007, p. 180)) and so

$$\begin{aligned}\int_{[0,1]\times(1,\infty)}\theta\frac{1}{k(n)}\sum_{i=1}^{n}\epsilon_{(\Theta_i,R_i/R_{(k(n))})}(d\theta,dr)&=\frac{1}{k(n)}\sum_{i=1}^{n}\Theta_i\mathbf{1}_{[R_i>R_{(k(n))}]}\\ &=\frac{1}{k(n)}\sum_{i=1}^{k(n)}\Theta_i^*\Rightarrow\int_{[0,1]\times(1,\infty)}\theta\epsilon_{\theta_0}\times\nu_\alpha(d\theta,dr)=\theta_0.\end{aligned}$$

Thus the denominator of (S5.16) is also $O_p(1)$. A successful comparison in (S5.15) has the difference converging to 0 and so it remains to show

$$\frac{1}{\sqrt{k(n)}}\Big|k(n)\theta_0-\sum_{i=1}^{k(n)}\Theta_i^*\Big|\Rightarrow 0. \tag{S5.17}$$

Since under $H_0^{(2)}$,

$$\frac{1}{\sqrt{k(n)}}\left|\sum_{i=1}^{k(n)}(\Theta_i^*-\theta_0)\right|=\frac{1}{\sqrt{k(n)}}\left|\sum_{i=1}^{n}(\Theta_i-\theta_0)\mathbf{1}_{\{R_i\ge R_{(k(n))}\}}\right|$$

$$\leq \frac{1}{\sqrt{k(n)}} \sum_{i=1}^{n} \frac{d(\boldsymbol{Z}_i, \mathbb{C}_{a,b})}{R_{(k(n))}} \mathbf{1}_{\{R_i \geq R_{(k(n))}\}} \xrightarrow{p} 0,$$

which can be seen as in the proof of (S4.2) by replacement of $R_{(k(n))}$ by $b(n/k(n))$ at the cost of $1 \pm \epsilon$ for any $\epsilon > 0$ with high probability. This confirms (S5.17) and thus proves convergence to 0 in (S5.15). In turn, this coupled with (S5.13) completes the proof.

Details for Step 2 in the proof of Theorem 4.1

The proof of (S5.13) requires justifying the application of the mapping

$$x \mapsto \int_1^\infty x(s) \frac{\mathrm{d}s}{s},$$

and applying this mapping to (S5.10) leads to

$$\frac{\sqrt{k(n)}}{\theta_0} \left(\frac{1}{k(n)} \sum_{i=1}^{k(n)} \Theta_i^* \log \frac{R_{(i)}}{R_{(k(n))}} - \frac{\theta_0}{\alpha} \left(\frac{R_{(k(n))}}{b(n/k(n))} \right)^{-\alpha} \right) \Rightarrow \int_1^\infty \frac{W(y^{-\alpha})}{y} \mathrm{d}y. \tag{S5.18}$$

For M large, applying the map

$$x \mapsto \int_1^M \frac{x(s)}{s} \mathrm{d}s$$

to (S5.10) gives

$$\frac{\sqrt{k(n)}}{\theta_0} \left(\int_1^M \frac{1}{k(n)} \sum_{i=1}^{n} \Theta_i \epsilon_{R_i/R_{(k(n))}}(s,\infty) \frac{\mathrm{d}s}{s} - \left(\frac{R_{(k(n))}}{b(n/k(n))} \right)^{-\alpha} \theta_0 \int_1^M s^{-\alpha-1} \mathrm{d}s \right)$$
$$\Rightarrow \int_1^M W(s^{-\alpha}) \frac{\mathrm{d}s}{s}.$$

As $M \to \infty$,

$$\int_1^M W(s^{-\alpha}) \frac{\mathrm{d}s}{s} \to \int_1^\infty \frac{W(s^{-\alpha})}{s} \mathrm{d}s.$$

Hence, it remains to verify that for any $\delta > 0$,

$$\lim_{M\to\infty}\limsup_{n\to\infty}\mathbb{P}\left(\sqrt{k(n)}\left|\int_M^\infty \frac{1}{k(n)}\sum_{i=1}^n \Theta_i\epsilon_{R_i/R_{(k(n))}}(s,\infty)\frac{\mathrm{d}s}{s} -\theta_0\left(\frac{R_{(k(n))}}{b(n/k(n))}\right)^{-\alpha}\int_M^\infty s^{-\alpha-1}\mathrm{d}s\right| > \delta\right) = 0 \quad \text{(S5.19)}$$

Rewrite the probability in (S5.19) as

$$\begin{aligned}
&\mathbb{P}\left(\sqrt{k(n)}\left|\int_M^\infty \left(\frac{1}{k(n)}\sum_{i=1}^n \Theta_i\epsilon_{R_i/R_{(k(n))}}(s,\infty) - \theta_0\left(\frac{R_{(k(n))}}{b(n/k(n))}s\right)^{-\alpha}\right)\frac{\mathrm{d}s}{s}\right| > \delta\right)\\
&\le \mathbb{P}\left(\sqrt{k(n)}\int_M^\infty \left|\frac{1}{k(n)}\sum_{i=1}^n \Theta_i\epsilon_{R_i/R_{(k(n))}}(s,\infty) - \theta_0\left(\frac{R_{(k(n))}}{b(n/k(n))}s\right)^{-\alpha}\right|\frac{\mathrm{d}s}{s} > \delta\right)\\
&= \mathbb{P}\left(\sqrt{k(n)}\int_{MR_{(k(n))}/b(n/k(n))}^\infty \left|\frac{1}{k(n)}\sum_{i=1}^n \Theta_i\epsilon_{R_i/b(n/k(n))}(s,\infty) - \theta_0 s^{-\alpha}\right|\frac{\mathrm{d}s}{s} > \delta\right)\\
&\le \mathbb{P}\left(\sqrt{k(n)}\int_{M(1-\eta)}^\infty \left|\frac{1}{k(n)}\sum_{i=1}^n \Theta_i\epsilon_{R_i/b(n/k(n))}(s,\infty) - \theta_0 s^{-\alpha}\right|\frac{\mathrm{d}s}{s} > \delta\right)\\
&\qquad + \mathbb{P}\left(\left|\frac{R_{(k(n))}}{b(n/k(n))} - 1\right| > \eta\right),
\end{aligned}$$

for $\eta > 0$. Since $R_{(k(n))}/b(n/k(n)) \xrightarrow{p} 1$, it suffices to consider

$$\begin{aligned}
&\mathbb{P}\left(\sqrt{k(n)}\int_{M(1-\eta)}^\infty \left|\frac{1}{k(n)}\sum_{i=1}^n \Theta_i\epsilon_{R_i/b(n/k(n))}(s,\infty) - \theta_0 s^{-\alpha}\right|\frac{\mathrm{d}s}{s} > \delta\right)\\
&\le \frac{k(n)}{\delta^2}\mathbb{E}\left[\int_{M(1-\eta)}^\infty \left(\frac{1}{k(n)}\sum_{i=1}^n \Theta_i\epsilon_{R_i/b(n/k(n))}(s,\infty) - \theta_0 s^{-\alpha}\right)^2\frac{\mathrm{d}s}{s}\right].
\end{aligned}$$

Write what is inside the square by centering the random term to get

$$\begin{aligned}
&\frac{1}{k(n)}\sum_{i=1}^n \Theta_i\epsilon_{R_i/b(n/k(n))}(s,\infty) - \frac{n}{k(n)}\mathbb{E}\left(\Theta_1\mathbf{1}_{\{R_1>b(n/k(n))s\}}\right)\\
&\quad + \frac{n}{k(n)}\mathbb{E}\left(\Theta_1\mathbf{1}_{\{R_1>b(n/k(n))s\}}\right) - \theta_0 s^{-\alpha}
\end{aligned}$$

and we get the probability in (S5.19) bounded by

$$\le \frac{k(n)}{\delta^2}\int_{M(1-\eta)}^\infty \frac{n}{k(n)^2}\mathrm{Var}\left(\Theta_1\mathbf{1}_{\{R_1>b(n/k(n))s\}}\right)\frac{\mathrm{d}s}{s}$$

$$+\frac{k(n)}{\delta^2}\int_{M(1-\eta)}^{\infty}\left(\frac{n}{k(n)}\mathbb{E}\left(\Theta_1\mathbf{1}_{\{R_1>b(n/k(n))s\}}\right)-\theta_0 s^{-\alpha}\right)^2\frac{\mathrm{d}s}{s}$$
$$=: I_n + II_n$$

Since $\Theta_1 \le 1$ a.s., the term I_n is bounded by

$$\frac{1}{\delta^2}\int_{M(1-\eta)}^{\infty}\frac{n}{k(n)}\mathbb{E}\left(\Theta_1\mathbf{1}_{\{R_1>b(n/k(n))s\}}\right)^2\frac{\mathrm{d}s}{s}\le\frac{1}{\delta^2}\int_{M(1-\eta)}^{\infty}\frac{n}{k(n)}\mathbb{P}(\{R_1>b(n/k(n))s)\frac{\mathrm{d}s}{s}.$$

By Karamata's theorem, the right side converges as $n\to\infty$ to

$$\frac{1}{\delta^2}\int_{M(1-\eta)}^{\infty}s^{-\alpha-1}\mathrm{d}s=\frac{1}{\alpha\delta^2}(M(1-\eta))^{-\alpha}\overset{M\to\infty}{\longrightarrow}0.$$

For II_n, with $v(t)=\mathbb{E}(\Theta_1\mathbf{1}_{\{R_1>t\}})$, we notice that

$$\begin{aligned}
&\frac{\theta_0^2}{A^2(n/k(n))}\left(\frac{n}{k(n)\theta_0}\mathbb{E}\left(\Theta_1\mathbf{1}_{\{R_1>b(n/k(n))s\}}\right)-s^{-\alpha}\right)^2\\
&=\frac{\theta_0^2}{A^2(n/k(n))}\left(\frac{n}{k(n)\theta_0}\mathbb{E}\left(\Theta_1\mathbf{1}_{\{R_1>b(n/k(n))s\}}\right)-\frac{v(b(n/k(n))s)}{v(b(n/k(n)))}+\frac{v(b(n/k(n))s)}{v(b(n/k(n)))}-s^{-\alpha}\right)^2\\
&\le\left|\frac{\theta_0}{A(n/k(n))}\left(\frac{n}{k(n)\theta_0}\mathbb{E}\left(\Theta_1\mathbf{1}_{\{R_1>b(n/k(n))s\}}\right)-\frac{v(b(n/k(n))s)}{v(b(n/k(n)))}\right)\right|^2\\
&\quad+\left|\frac{\theta_0}{A(n/k(n))}\left(\frac{v(b(n/k(n))s)}{v(b(n/k(n)))}-s^{-\alpha}\right)\right|^2\\
&\quad+2\left|\frac{\theta_0}{A(n/k(n))}\left(\frac{n}{k(n)\theta_0}\mathbb{E}\left(\Theta_1\mathbf{1}_{\{R_1>b(n/k(n))s\}}\right)-\frac{v(b(n/k(n))s)}{v(b(n/k(n)))}\right)\right|\\
&\qquad\times\left|\frac{\theta_0}{A(n/k(n))}\left(\frac{v(b(n/k(n))s)}{v(b(n/k(n)))}-s^{-\alpha}\right)\right|.
\end{aligned}$$

By (de Haan and Ferreira, 2006, Theorem 2.3.9), for any $\epsilon>0$, $\delta\in(0,\alpha(1+\rho))$, there exists $A_0(n/k(n))\sim A(n/k(n))$ as $n\to\infty$ and $n_0\equiv n_0(\epsilon,\delta)$ such that for all $b(n/k(n)), b(n/k(n))s\ge n_0$,

$$\left|\frac{1}{A_0(n/k(n))}\left(\frac{v(b(n/k(n))s)}{v(b(n/k(n)))}-s^{-\alpha}\right)-s^{-\alpha}\frac{1-s^{-\alpha\rho}}{\alpha\rho}\right|\le\epsilon s^{-\alpha(1+\rho)}\max\{s^{-\delta},s^{\delta}\},$$

so that

$$\begin{aligned}&\left|\frac{1}{A_0(n/k(n))}\left(\frac{v(b(n/k(n))s)}{v(b(n/k(n)))}-s^{-\alpha}\right)\right|\\&\leq s^{-\alpha}\left(\left|\frac{1-s^{-\alpha\rho}}{\alpha\rho}\right|+\epsilon s^{-\alpha\rho}\max\{s^{-\delta},s^{\delta}\}\right).\end{aligned}$$

Hence,

$$\begin{aligned}&\int_{M(1-\eta)}^{\infty}\left|\frac{\theta_0}{A_0(n/k(n))}\left(\frac{v(b(n/k(n))s)}{v(b(n/k(n)))}-s^{-\alpha}\right)\right|^2\frac{\mathrm{d}s}{s}\\&\qquad\leq\int_{M(1-\eta)}^{\infty}s^{-2\alpha}\left(\left|\frac{1-s^{-\alpha\rho}}{\alpha\rho}\right|+\epsilon s^{-\alpha\rho}\max\{s^{-\delta},s^{\delta}\}\right)^2\frac{\mathrm{d}s}{s}<\infty,\end{aligned}$$

which further implies

$$\int_{M(1-\eta)}^{\infty}\left|\frac{\theta_0}{A(n/k(n))}\left(\frac{v(b(n/k(n))s)}{v(b(n/k(n)))}-s^{-\alpha}\right)\right|^2\frac{\mathrm{d}s}{s}<\infty,$$

as $A_0(n/k(n))\sim A(n/k(n))$. In addition, since

$$\begin{aligned}&\frac{\theta_0}{A(n/k(n))}\left(\frac{n}{k(n)\theta_0}\mathbb{E}\left(\Theta_1\mathbf{1}_{\{R_1>b(n/k(n))s\}}\right)-\frac{v(b(n/k(n))s)}{v(b(n/k(n)))}\right)\\&\qquad=\frac{v(b(n/k(n))s)}{v(b(n/k(n)))\theta_0}\frac{1}{A(n/k(n))}\left(\frac{n}{k(n)}v(b(n/k(n)))-\theta_0\right),\end{aligned}$$

then

$$\begin{aligned}&\int_{M(1-\eta)}^{\infty}\left|\frac{\theta_0}{A(n/k(n))}\left(\frac{n}{k(n)\theta_0}\mathbb{E}\left(\Theta_1\mathbf{1}_{\{R_1>b(n/k(n))s\}}\right)-\frac{v(b(n/k(n))s)}{v(b(n/k(n)))}\right)\right|^2\frac{\mathrm{d}s}{s}\\&\qquad=\left|\frac{1}{A(n/k(n))}\left(\frac{n}{k(n)}v(b(n/k(n)))-\theta_0\right)\right|^2\int_{M(1-\eta)}^{\infty}\frac{v^2(b(n/k(n))s)}{\theta_0^2v^2(b(n/k(n)))}\frac{\mathrm{d}s}{s}\\&\qquad\to\left(\frac{1}{\theta_0}\int\int_{(1,\infty)\times[0,1]}\theta\chi(\mathrm{d}r,\mathrm{d}\theta)\right)^2\frac{(M(1-\eta))^{-2\alpha}}{2\alpha}<\infty.\end{aligned}$$

Since under the condition in Eq. (3.18), the intermediate sequence $\{k(n)\}$ also satisfies $\sqrt{k(n)}A(n/k(n))\to 0$ as $n\to\infty$. We then see that

$$k(n)\int_{M(1-\eta)}^{\infty}\left|\left(\frac{n}{k(n)}\mathbb{E}\left(\Theta_1\mathbf{1}_{\{R_1>b(n/k(n))s\}}\right)-\theta_0\frac{v(b(n/k(n))s)}{v(b(n/k(n)))}\right)\right|^2\frac{\mathrm{d}s}{s}$$

$$=(\sqrt{k(n)}A(n/k(n)))^2$$
$$\times \int_{M(1-\eta)}^{\infty} \left| \frac{\theta_0}{A(n/k(n))} \left(\frac{n}{k(n)\theta_0} \mathbb{E}\left(\Theta_1 \mathbf{1}_{\{R_1 > b(n/k(n))s\}}\right) - \frac{v(b(n/k(n))s)}{v(b(n/k(n)))} \right) \right|^2 \frac{\mathrm{d}s}{s} \to 0,$$

and similarly,

$$k(n) \int_{M(1-\eta)}^{\infty} \left| \theta_0 \left(\frac{v(b(n/k(n))s)}{v(b(n/k(n)))} - s^{-\alpha} \right) \right|^2 \frac{\mathrm{d}s}{s} \to 0.$$

So we conclude that $II_n \to 0$. This justifies (S5.19), thus completing the proof of (S5.18).

Recall (S5.11), and applying the mapping

$$(x, y) \mapsto \left(\int_1^{\infty} x(s) \frac{\mathrm{d}s}{s}, y \right)$$

to (S5.11) gives

$$\frac{\sqrt{k(n)}}{\theta_0} \left(\frac{1}{k(n)} \sum_{i=1}^{k(n)} \Theta_i^* \log \frac{R_{(i)}}{R_{(k(n))}} - \frac{\theta_0}{\alpha} \right) \Rightarrow \int_1^{\infty} \frac{W(y^{-\alpha})}{y} \mathrm{d}y - \frac{1}{\alpha} W(1). \tag{S5.20}$$

S6 Proof of Theorem 4.2

Proceed by steps.

(1) First, we need to show that in $D(0, 1]$,

$$\frac{\sqrt{k(n)}}{(1 + \sigma^2/\mu^2)^{1/2}} \left(\frac{1}{\mu k(n)} \sum_{i=1}^{n} \Theta_i \epsilon_{R_i/b(n/k(n))}(t^{-1/\alpha}, \infty) - t \right) \Rightarrow W(t), \tag{S6.21}$$

where $W(\cdot)$ is a standard Brownian motion. This is obtained by employing the functional central limit theorem given in Theorem S2.1 (details deferred to Section S6) to get in $D(0,1]$,

$$\frac{(\mu^2+\sigma^2)^{-1/2}}{\sqrt{k(n)}}\sum_{i=1}^{n}\left(\Theta_i\epsilon_{R_i/b(n/k(n))}(t^{-1/\alpha},\infty)-\mathbb{E}\big(\Theta_i\epsilon_{R_i/b(n/k(n))}(t^{-1/\alpha},\infty)\big)\right)\Rightarrow W(t). \tag{S6.22}$$

Note that by 2RV using Eq. (2.9) or Eq. (2.10) plus the marginalized version for the distribution of R_1, we have

$$\sqrt{k(n)}\left(\frac{n}{k(n)}\mathbb{E}\left(\Theta_1\mathbf{1}_{\{R_1>b(n/k(n))y\}}\right)-\mu\frac{n}{k(n)}\mathbb{P}(R_1>b(n/k(n))y)\right)\to 0 \tag{S6.23}$$

locally uniformly for $y>0$ and for $\{k(n)\}$ satisfying Eq. (3.18). Combining (S6.23) with (S6.22) then completes the proof of (S6.21).

(2) Applying the composition map $(x(t),c)\mapsto x(ct)$ from $D(0,\infty)\times(0,\infty)\mapsto D(0,\infty)$, we get in $D(0,\infty)$,

$$\frac{\sqrt{k(n)}}{(1+\sigma^2/\mu^2)^{1/2}}\left(\frac{1}{\mu k(n)}\sum_{i=1}^{n}\Theta_i\epsilon_{R_i/R_{(k(n))}}(y,\infty)-\left(y\frac{R_{(k(n))}}{b(n/k(n))}\right)^{-\alpha}\right)\Rightarrow W(y^{-\alpha}). \tag{S6.24}$$

Repeating a similar argument as in Step 2 of the proof for Theorem 4.1, we are able to justify the application of

$$x\mapsto\int_1^\infty\frac{x(s)}{s}\mathrm{d}s,$$

which further leads to

$$(\mu^2+\sigma^2)^{-1/2}\sqrt{k(n)}\left(\frac{1}{\mu k(n)}\sum_{i=1}^{k(n)}\Theta_i^*\log\frac{R_{(i)}}{R_{(k(n))}}-\frac{1}{\alpha}\left(\frac{R_{(k(n))}}{b(n/k(n))}\right)^{-\alpha}\right)$$
$$\Rightarrow\frac{1}{\alpha}\int_0^1\frac{W(s)}{s}\mathrm{d}s. \tag{S6.25}$$

(3) We are left to justify the convergence of

$$\begin{aligned}
&\sqrt{k(n)}(T_n - 1/\alpha) - \sqrt{k(n)}\left(\frac{1}{\mu k(n)}\sum_{i=1}^{k(n)} \Theta_i^* \log \frac{R_{(i)}}{R_{(k(n))}} - \frac{1}{\alpha}\left(\frac{R_{(k(n))}}{b(n/k(n))}\right)^{-\alpha}\right) \\
&\qquad = \sqrt{k(n)}\left(\left(\frac{1}{k(n)}\sum_{i=1}^{k(n)}\Theta_i^*\right)^{-1} - \frac{1}{\mu}\right)\frac{1}{k(n)}\sum_{i=1}^{k(n)}\Theta_i^* \log\frac{R_{(i)}}{R_{(k(n))}} \\
&\qquad + \frac{\sqrt{k(n)}}{\alpha}\left(\left(\frac{R_{(k(n))}}{b(n/k(n))}\right)^{-\alpha} - 1\right).
\end{aligned}$$

Note that by (S6.24),

$$\frac{1}{k(n)}\sum_{i=1}^{k(n)}\Theta_i^* = \frac{1}{k(n)}\sum_{i=1}^{n}\Theta_i \epsilon_{R_i/R_{(k(n))}}(1,\infty) \xrightarrow{p} \mu,$$

and the convergence in (S6.25) gives

$$\frac{1}{k(n)}\sum_{i=1}^{k(n)}\Theta_i^* \log\frac{R_{(i)}}{R_{(k(n))}} \xrightarrow{p} \frac{\mu}{\alpha}.$$

Therefore, it suffices to consider the convergence of

$$\frac{\sqrt{k(n)}}{\alpha}\left(\left(\frac{R_{(k(n))}}{b(n/k(n))}\right)^{-\alpha} - \frac{1}{\mu k(n)}\sum_{i=1}^{k(n)}\Theta_i^*\right). \tag{S6.26}$$

To prove the convergence of (S6.26), we first use Vervaat's inversion (Vervaat (1972); de Haan and Ferreira (2006); Resnick (2007)) to obtain the convergence of the inverse of the function

$$\eta_n(\cdot) := \frac{1}{\mu k(n)}\sum_{i=1}^{n}\Theta_i \epsilon_{R_i/b(n/k(n))}\big((\cdot)^{-1/\alpha},\infty\big).$$

Note that

$$\begin{aligned}
\eta_n^{\leftarrow}(t) &= \inf\left\{s : \frac{1}{k(n)}\sum_{i=1}^{n}\Theta_i \epsilon_{R_i/b(n/k(n))}\big(s^{-1/\alpha},\infty\big) \geq \mu t\right\} \\
&= \left(\sup\left\{y : \frac{1}{k(n)}\sum_{i=1}^{n}\Theta_i \epsilon_{R_i/b(n/k(n))}\big(y,\infty\big) \geq \mu t\right\}\right)^{-\alpha}.
\end{aligned} \tag{S6.27}$$

Then with

$$m(t) := \inf\left\{ m \geq 1 : \frac{1}{k(n)} \sum_{i=1}^{n} \Theta_i \epsilon_{R_i/b(n/k(n))} \left(\frac{R_{(m)}}{b(n/k(n))}, \infty \right) \geq \mu t \right\},$$

the inverse function in (S6.27) becomes

$$\eta_n^{\leftarrow}(t) = \left(\frac{R_{(m(t))}}{b(n/k(n))} \right)^{-\alpha}.$$

Applying Vervaat's lemma (Vervaat (1972); de Haan and Ferreira (2006); Resnick (2007)) gives the joint convergence in $D(0,1] \times D(0,1]$:

$$\frac{\sqrt{k(n)}}{(1+\sigma^2/\mu^2)^{1/2}} \left(\eta_n(t) - t,\, \eta_n^{\leftarrow}(t) - t \right) \Rightarrow \left(W(t), -W(t) \right). \tag{S6.28}$$

Since for $t = \frac{1}{\mu k(n)} \sum_{i=1}^{k(n)} \Theta_i^*$,

$$m\left(\frac{1}{\mu k(n)} \sum_{i=1}^{k(n)} \Theta_i^* \right)$$
$$= \inf\left\{ m \geq 1 : \frac{1}{k(n)} \sum_{i=1}^{n} \Theta_i \epsilon_{R_i/b(n/k(n))} \left(\frac{R_{(m)}}{b(n/k(n))}, \infty \right) \geq \frac{1}{k(n)} \sum_{i=1}^{k(n)} \Theta_i^* \right\} = k(n),$$

then (S6.28) gives

$$\sqrt{k(n)} \left(\left(\frac{R_{(k(n))}}{b(n/k(n))} \right)^{-\alpha} - \frac{1}{\mu k(n)} \sum_{i=1}^{k(n)} \Theta_i^* \right) \Rightarrow -(1+\sigma^2/\mu^2)^{1/2} W(1). \tag{S6.29}$$

Combining (S6.29) with (S6.25) shows that

$$\sqrt{k(n)}\,(T_n - 1/\alpha) \Rightarrow \frac{(1+\sigma^2/\mu^2)^{1/2}}{\alpha} \left(\int_0^1 \frac{W(s)}{s} \mathrm{d}s - W(1) \right),$$

thus verifying Eq. (4.25).

Verify conditions for FCLT.

To align with the statement in Theorem S2.1, we define

$$f_{n,i}(t) := \frac{(\mu^2+\sigma^2)^{-1/2}}{\sqrt{k(n)}}\Theta_i\epsilon_{R_i/b(n/k(n))}(t^{-1/\alpha},\infty), \qquad t\in(0,1],$$

and the envelope function

$$F_{n,i} := \frac{(\mu^2+\sigma^2)^{-1/2}}{\sqrt{k(n)}}\epsilon_{R_i/b(n/k(n))}(1,\infty).$$

By Definition 7.9 of Pollard (1990), we see that $\{f_{n,i}\}$ are manageable. Also, with

$$X_n(t) = \sum_{i=1}^n \left(f_{n,i}(t) - \mathbb{E}[f_{n,i}(t)]\right),$$

we have

$$\begin{aligned}\mathbb{E}\left(X_n(t)X_n(s)\right) &= \frac{(\mu^2+\sigma^2)^{-1}}{k(n)}\sum_{i=1}^n \mathbb{E}\left(\Theta_i^2\epsilon_{R_i/b(n/k(n))}\big((t\wedge s)^{-1/\alpha},\infty\big)\right)\\ &- \frac{(\mu^2+\sigma^2)^{-1}}{k(n)}\sum_{i=1}^n \mathbb{E}\left(\Theta_i\epsilon_{R_i/b(n/k(n))}\big(t^{-1/\alpha},\infty\big)\right)\mathbb{E}\left(\Theta_i\epsilon_{R_i/b(n/k(n))}\big(s^{-1/\alpha},\infty\big)\right)\\ &\longrightarrow t\wedge s, \qquad n\to\infty.\end{aligned}$$

For the envelope function $F_{n,i}$, we have

$$\limsup_{n\to\infty}\sum_{i=1}^n \mathbb{E}(F_{n,i}^2) \to (\mu^2+\sigma^2)^{-1},$$

and for $\delta>0$,

$$\sum_{i=1}^n \mathbb{E}(F_{n,i}^{2+\delta}) = (\mu^2+\sigma^2)^{-(1+\delta/2)}\frac{n}{k(n)^{1+\delta/2}}\mathbb{P}\left(\frac{R_1}{b(n/k(n))}>1\right)\to 0,$$

which further implies for each $\eta>0$,

$$\sum_{i=1}^n \mathbb{E}(F_{n,i}^2\mathbf{1}_{\{F_{n,i}>\eta\}}) \to 0.$$

Assume $t_1 > t_2$, then for

$$\rho_n(t_1,t_2) := \left(\sum_{i=1}^{n} \mathbb{E}\left(f_{n,i}(t_1) - f_{n,i}(t_2)\right)^2\right)^{1/2},$$

we have

$$\begin{aligned}\rho_n(t_1,t_2) &= (\mu^2+\sigma^2)^{-1/2}\left(\frac{n}{k(n)}\mathbb{E}\left(\Theta_1\epsilon_{R_1/b(n/k(n))}(t_1^{-1/\alpha}, t_2^{-1/\alpha})\right)\right)^{1/2}\\ &\longrightarrow (\mu^2+\sigma^2)^{-1/2}(t_1-t_2).\end{aligned}$$

Therefore, all conditions in Theorem S2.1 are satisfied, which gives the conclusion that in $D(0,1]$,

$$X_n(t) = \frac{(\mu^2+\sigma^2)^{-1/2}}{\sqrt{k(n)}}\sum_{i=1}^{n}\left(\Theta_i\epsilon_{R_i/b(n/k(n))}(t^{-1/\alpha},\infty) - \mathbb{E}\big(\Theta_i\epsilon_{R_i/b(n/k(n))}(t^{-1/\alpha},\infty)\big)\right) \Rightarrow W(t).$$

S7 Proof of Theorem 5.1

To shorten the proof and ease notation we make the simplifying assumption that we know $b = 1$. Define,

$$\mathbb{C}_s = \mathbb{C}_{s,1} = \{(x,y)\in R_+^2 : s \le x/(x+y)\},$$

$$d_s\big((x,y)\big) = d_{s,1}\big((x,y)\big) = d\big((x,y),\mathbb{C}_s\big) = \left\{y - (s^{-1}-1)x\right\}^+, \qquad 0 \le s \le 1.$$

$$g_n(s) = g_n(s,1) = (1-s) + \frac{1}{\sqrt{k(n)}}\sum_{i=1}^{k(n)} \frac{d_s\big(\boldsymbol{Z}_i^*\big)}{R_{(k(n))}}\log\frac{R_{(i)}}{R_{(k(n))}}.$$

Note that $g_n(\cdot)$ is concave in s. We prove Theorem 5.1 by showing that for some $\epsilon > 0$, and $\mathcal{I}_\epsilon := [0, a-\epsilon] \cup [a+\epsilon, 1]$,

$$\mathbb{P}\left(\inf_{s\in\mathcal{I}_\epsilon}(g_n(s) - g_n(a)) > 2\epsilon\right) \to 1, \quad (n\to\infty). \tag{S7.30}$$

Since $d_s\big((x,y)\big)$ is increasing in s, Theorem 3.1 ensures that (see Eq. (3.20))

$$\sup_{s\in[0,a]} \frac{1}{\sqrt{k(n)}} \sum_{i=1}^{k(n)} \frac{d_s\big(\boldsymbol{Z}_i^*\big)}{R_{(k(n))}} \log \frac{R_{(i)}}{R_{(k(n))}} \xrightarrow{p} 0,$$

and therefore,

$$\lim_{n\to\infty} \mathbb{P}\left(\inf_{s\le a-\epsilon}(g_n(s)-g_n(a)) > 2\epsilon\right) = 1. \tag{S7.31}$$

If $s \ge a+\epsilon$, replace division by $\sqrt{k(n)}$ with division by $k(n)$ and the definitions of (R_i,Θ_i) give

$$\frac{1}{k(n)} \sum_{i=1}^{k(n)} \frac{d_s\big(\boldsymbol{Z}_i^*\big)}{R_{(k(n))}} \log \frac{R_{(i)}}{R_{(k(n))}} = \frac{1}{k(n)} \sum_{i=1}^{k(n)} \big\{1-s^{-1}\Theta_i^*\big\}^+ \frac{R_{(i)}}{R_{(k(n))}} \log \frac{R_{(i)}}{R_{(k(n))}},$$

and since $R_{(i)} \ge R_{(k(n))}$, this is bounded below by

$$\ge \frac{1}{k(n)} \sum_{i=1}^{k(n)} \big\{1-s^{-1}\Theta_i^*\big\}^+ \log \frac{R_{(i)}}{R_{(k(n))}}. \tag{S7.32}$$

We show this converges in probability to a positive constant $L(a,s) > 0$ and thus

$$\frac{1}{\sqrt{k(n)}} \sum_{i=1}^{k(n)} \frac{d_s\big(\boldsymbol{Z}_i^*\big)}{R_{(k(n))}} \log \frac{R_{(i)}}{R_{(k(n))}} \xrightarrow{p} \infty. \tag{S7.33}$$

Combining (S7.33) with (S7.31) completes the proof of (S7.30).

Returning to the expression in (S7.32), write it as

$$\iint_{\{(r,\theta):r>1,\theta\in[0,1]\}} \Big(1-\frac{\theta}{s}\Big)^+ \log r \frac{1}{k} \sum_{i=1}^{k(n)} \epsilon_{(\Theta_i^*, R_{(i)}/R_{(k)})}(d\theta, dr). \tag{S7.34}$$

On $[0,1]\times(1,\infty)$,

$$\frac{1}{k} \sum_{i=1}^{k(n)} \epsilon_{(\Theta_i^*, R_{(i)}/R_{(k)})}(d\theta, dr) \to S(d\theta)\nu_\alpha(dr)$$

and the right side is a probability measure on $[0,1]\times(1,\infty)$, so using familiar weak convergence arguments in (S7.34) we get convergence to

$$L(a,s) := \iint_{\{(r,\theta):r>1,\theta\in[0,1]\}} \Big(1-\frac{\theta}{s}\Big)^{+} \log r S(d\theta)\nu_\alpha(dr)$$

and after some Fubini justified manipulations this is

$$=\frac{1}{\alpha}\int_a^s \Big(1-\frac{\theta}{s}\Big) S(d\theta).$$

Remember $s > a+\epsilon$ and we verify $L(a,s)$ is positive. If not, $L(a,s)=0$ and $1-\theta/s=0$ or $\theta=s$ for almost all (with respect to $S(\cdot)$) $\theta\in[a,,a+\epsilon]$ and this means $S[a,a+\epsilon]=0$, thus contradicting a being in the support of $S(\cdot)$. So $L(a,s)>0$.

References

de Haan, L. and A. Ferreira (2006). *Extreme Value Theory: An Introduction.* New York: Springer-Verlag.

Pollard, D. (1990). *Empirical Processes: Theory and Applications.* NSF-CBMS Regional Conference Series in Probability and Statistics.

Resnick, S. (2007). *Heavy Tail Phenomena: Probabilistic and Statistical Modeling.* Springer Series in Operations Research and Financial Engineering. New York: Springer-Verlag. ISBN: 0-387-24272-4.

Vervaat, W. (1972). Functional central limit theorems for processes with positive drift and their inverses. *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete 23*, 245–253.