Refining the Central Limit Theorem Approximation via Extreme Value Theory

Ulrich K. Müller*
Economics Department
Princeton University

September 2018

Abstract

We suggest approximating the distribution of the sum of independent and identically distributed random variables with a Pareto-like tail by combining extreme value approximations for the largest summands with a normal approximation for the sum of the smaller summands. If the tail is well approximated by a Pareto density, then this new approximation has substantially smaller error rates compared to the usual normal approximation for underlying distributions with finite variance and less than three moments. It can also provide an accurate approximation for some infinite variance distributions.

AMS 2000 subject classification: 60F05, 60G70, 60E07

Key words and phrases: Regular variation, rates of convergence

*Funding via NSF grant SES-1627660 is gratefully acknowledged.

1 Introduction

Consider approximations to the distribution of the sum $S_n = \sum_{i=1}^n X_i$ of independent mean-zero random variables X_i with distribution function F. If $\sigma_0^2 = \int x^2 dF(x)$ exists, then $n^{-1/2}S_n$ is asymptotically normal by the central limit theorem. The quality of this approximation is poor if $\max_{i\leq n}|X_i|$ is not much smaller than $n^{1/2}$, since then a single non-normal random variable has non-negligible influence on $n^{-1/2}S_n$. Extreme value theory provides large sample approximations to the behavior of the largest observations, suggesting that it may be fruitfully employed in the derivation of better approximations to the distribution of S_n.

For simplicity, consider the case where F has a light left tail and a heavy right tail. Specifically, assume $\int_{-\infty}^0 |x|^3 dF(x) < \infty$ and

$$\lim_{x\to\infty} \frac{1-F(x)}{x^{-1/\xi}} = \omega^{1/\xi}, \omega > 0 \tag{1}$$

for $1/3 < \xi < 1$, so that the right tail of F is approximately Pareto with shape parameter $1/\xi$ and scale parameter ω. Let $X_{i:n}$ be the order statistics. For a given sequence $k = k(n)$, $1 \leq k < n$, split S_n into two pieces

$$S_n = \sum_{i=1}^{n-k} X_{i:n} + \sum_{i=1}^{k} X_{n-i+1:n}. \tag{2}$$

Note that conditional on the $n-k$th order statistic $T_n = X_{n-k+1:n}$, $\sum_{i=1}^{n-k} X_{i:n}$ has the same distribution as $\sum_{i=1}^{n-k} \tilde{X}_i$, where $\tilde{X}_i$ are i.i.d. from the truncated distribution $\tilde{F}_{T_n}(x)$ with $\tilde{F}_t(x) = F(x)/F(t)$ for $x \leq t$ and $\tilde{F}_t(x) = 1$ otherwise. Let $\mu(t)$ and $\sigma^2(t)$ be the mean and variance of $\tilde{F}_t$. Since $\tilde{F}_{T_n}$ is less skewed than F, one would expect the distributional approximation (denoted by "$\overset{a}{\sim}$") of the central limit theorem,

$$\sum_{i=1}^{n-k} X_{i:n}|T_n \overset{a}{\sim} (n-k)\mu(T_n) + (n-k)^{1/2}\sigma(T_n)Z \quad \text{for} \quad Z \sim \mathcal{N}(0,1) \tag{3}$$

to be relatively accurate. At the same time, extreme value theory implies that under (1),

$$\sum_{i=1}^{k} X_{n-i+1:n} \overset{a}{\sim} n^{\xi}\omega \sum_{i=1}^{k} \Gamma_i^{-\xi} \text{ for } \Gamma_i = \sum_{j=1}^{i} E_j,\ E_j \sim \text{i.i.d. exponential.} \tag{4}$$

Combining (3) and (4) suggests

$$S_n \overset{a}{\sim} n^{1/2}\sigma(n^{\xi}\Gamma_k^{-\xi})Z + (n-k)\mu(n^{\xi}\Gamma_k^{-\xi}) + n^{\xi}\omega\sum_{i=1}^{k}\Gamma_i^{-\xi} \tag{5}$$

with Z independent of $(\Gamma_i)_{i=1}^k$.

If $\xi < 1/2$, the approximate Pareto tail (1) and $\mathbb{E}[X_1] = 0$ imply

$$\mu(x) \approx -\frac{\omega^{1/\xi}x^{1-1/\xi}}{(1-\xi)(1-(x/\omega)^{-1/\xi})}$$

and $\sigma^2(x) \approx \sigma_0^2 - \omega^{1/\xi}\frac{1}{1-2\xi}x^{2-1/\xi}$ for x large. From $(n-k)/(n-\Gamma_k) \overset{a}{\sim} 1$, this further yields

$$S_n \overset{a}{\sim} n^{1/2}\left(\sigma_0^2 - \frac{\omega^2}{1-2\xi}(\Gamma_k/n)^{1-2\xi}\right)^{1/2} Z - n^{\xi}\frac{\omega}{1-\xi}\Gamma_k^{1-\xi} + n^{\xi}\omega\sum_{i=1}^{k}\Gamma_i^{-\xi} \tag{6}$$

which depends on F only through the unconditional variance σ_0^2 and the two tail parameters (ω,ξ). Note that $\mathbb{E}[\Gamma_i^{-\xi}] = \Gamma(i-\xi)/\Gamma(i)$ and $\mathbb{E}[\Gamma_k^{1-\xi}] = \Gamma(1+k-\xi)/\Gamma(k) = (1-\xi)\sum_{i=1}^k \Gamma(i-\xi)/\Gamma(i)$, so the right-hand side of (6) is the sum of a mean-zero right skewed random variable, and a (dependent) random-scale mean-zero normal variable.

Theorem 1 below provides an upper bound on the convergence rate of the error in the approximation (6). The proof combines the Berry-Esseen bound for the central limit theorem approximation in (3) and the rate result in Corollary 5.5.5 of Reiss (1989) for the extreme value approximation in (4). If the tail of F is such that the approximation in (4) is accurate, then for both fixed and diverging k the error in (6) converges to zero faster than the error in the usual mean-zero normal approximation. The approximation (6) thus helps illuminate the nature and origin of the leading error terms in the first order normal approximation, as derived in Chapter 2 of Hall (1982), for such F. We also provide a characterization of the bound minimizing choice of k.

If $\xi > 1/2$, then the distribution of $n^{-\xi}S_n$ converges to a one-sided stable law with index ξ. An elegant argument by LePage, Woodroofe, and Zinn (1981) shows that this limiting law can be written as $\omega\sum_{i=1}^{\infty}\Gamma_i^{-\xi}$. The approximation (5) thus remains potentially accurate under $k \to \infty$ also for infinite variance distributions. To obtain a further approximation akin to (6), note that (1) implies $\sigma^2(\omega x) - \sigma^2(\omega y) \approx (\omega^2/\xi)\int_y^x t^{1-1/\xi}dt$ for large x, y. Let $u_n = (n/k)^{\xi}$. Then

$$S_n \overset{a}{\sim} n^{1/2}\left(\sigma^2(\omega u_n) + \frac{\omega^2}{\xi}\int_{u_n}^{(n/\Gamma_k)^{\xi}} y^{1-1/\xi}dy\right)^{1/2} Z - n^{\xi}\frac{\omega}{1-\xi}\Gamma_k^{1-\xi} + n^{\xi}\omega\sum_{i=1}^{k}\Gamma_i^{-\xi} \tag{7}$$

which depends on F only through the tail parameters (ω, ξ) and the sequence of truncated variances $\sigma^2(\omega u_n)$. The approximation (7) could also be applied to the case $\xi < 1/2$, so that one obtains a unifying approximation for values of ξ both smaller and larger than $1/2$. Indeed, for F mean-centered Pareto of index ξ, the results below imply that for suitable choice of $k \to \infty$, this approximation has an error that converges to zero much faster than the error from the first order approximation via the normal or non-normal stable limit for ξ close to $1/2$. The approach here thus also sheds light on the nature of the leading error terms of the non-normal stable limit, such as those derived by Christoph and Wolf (1992).

For $\xi > 1/2$, the idea of splitting up S_n as in (2) and to jointly analyze the asymptotic behavior of the pieces is already pursued in Csörgö, Haeusler, and Mason (1988). The contribution here is to derive error rates for resulting approximation to the distribution of the sum, especially for $1/3 < \xi < 1/2$, and to develop the additional approximation of the truncated mean and variance induced by the approximate Pareto tail.

The next section formalizes these arguments and discusses various forms of writing the variance term and the approximation for the case where both tails are heavy. Section 3 contains the proofs.

2 Assumptions and Main Results

The following condition imposes the right tail of F to be in the δ-neighborhood of the Pareto distribution with index ξ, as defined in Chapter 2 of Falk, Hüsler, and Reiss (2004).

Condition 1 *For some $x_0, \delta, \omega, L_F > 0$ and $1/3 < \xi < 1$, $F(x)$ admits a density for all $x \geq x_0$ of the form*

$$f(x) = (\omega\xi)^{-1}(x/\omega)^{-1/\xi-1}(1 + h(x)) \tag{8}$$

with $|h(x)| \leq L_F x^{-\delta/\xi}$ uniformly in $x \geq x_0$.

As discussed in Falk, Hüsler, and Reiss (2004), Condition 1 can be motivated by considering the remainder in the von Mises condition for extreme value theory. It is also closely related to the assumption that the tail of F is second order regularly varying, as studied by de Haan and Stadtmüller (1996) and de Haan and Resnick (1996). Many heavy-tailed distributions satisfy Condition 1: for the right tail of a student-t distribution with ν degrees of freedom, $\xi = 1/\nu$ and $\delta = 2\xi$, for the tail of a Fréchet or generalized extreme value distribution with parameter α, $\xi = 1/\alpha$ and $\delta = 1$, and for an exact Pareto tail, δ may be chosen arbitrarily large. In general, shifts of the distribution affect δ; for instance, a mean-centered Pareto distribution satisfies Condition 1 only for $\delta \leq \xi$. See Remark 4 below.

Not all heavy-tailed distributions in the domain of attraction of a Fréchet limit law satisfy Condition 1. A density of the form (8) with $h(x) = 1/\log(1+x)$, for example, does not. Under some additional regularity conditions on the von Mises remainder term, Theorem 3.2 of Falk and Marohn (1993) shows Condition 1 to be necessary to obtain an error rate of extreme value approximations of order $n^{-\delta}$ for $\delta > 0$. Roughly speaking, Condition 1 thus merely formalizes the assumption that extreme value theory provides accurate approximations.

We write C for a generic positive constant that does not depend on k or n, not necessarily the same in each instance it is used.

Theorem 1 *Under Condition 1,*

(a) for $1/3 < \xi < 1/2$

$$\sup_s \left| \mathbb{P}(n^{-1/2} S_n \leq s) - \mathbb{P}\left(n^{1/2} \left(\sigma_0^2 - \frac{\omega^2}{1-2\xi} (\Gamma_k/n)^{1-2\xi} \right)^{1/2} Z \right.\right.$$
$$\left.\left. - n^{\xi} \frac{\omega}{1-\xi} \Gamma_k^{1-\xi} + n^{\xi} \omega \sum_{i=1}^{k} \Gamma_i^{-\xi} \leq s n^{1/2} \right) \right| \leq C \cdot R(k,n,\xi,\delta)$$

(b) for $1/3 < \xi < 1$, $u_n = (n/k)^{\xi}$ *and* $a_n = (n \log n)^{-1/2}$ *for* $\xi = 1/2$ *and* $a_n = n^{-\max(\xi,1/2)}$ *otherwise,*

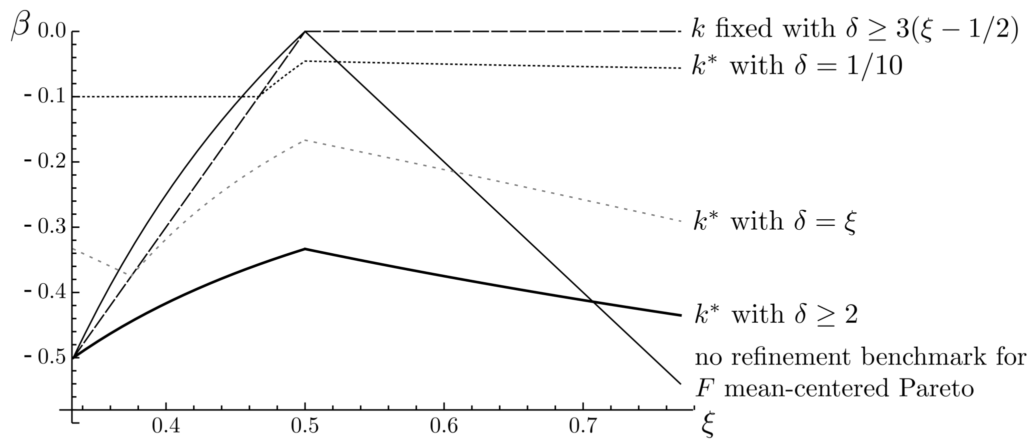
$$\sup_s \left| \mathbb{P}(a_n S_n \leq s) - \mathbb{P}\left(n^{1/2} \left(\sigma^2(\omega u_n) + \frac{\omega^2}{\xi} \int_{u_n}^{(n/\Gamma_k)^{\xi}} y^{1-1/\xi} dy \right)^{1/2} Z \right.\right.$$
$$\left.\left. - n^{\xi} \frac{\omega}{1-\xi} \Gamma_k^{1-\xi} + n^{\xi} \omega \sum_{i=1}^{k} \Gamma_i^{-\xi} \leq s/a_n \right) \right| \leq C \cdot R(k,n,\xi,\delta)$$

where

$$R(k,n,\xi,\delta) = \begin{cases} n^{-1/2}(n/k)^{3\xi-1} + (k/n)^{\delta} k^{1/2} + k/n & \text{for } 1/3 < \xi < 1/2 \\ k^{-\xi} + (k/n)^{\delta} k^{1/2} + k/n & \text{for } 1/2 \leq \xi < 1. \end{cases}$$

It is straightforward to characterize the rate for k which minimizes the bound $R(k,n,\xi,\delta)$. For two positive sequences a_n, b_n, write $a_n \asymp b_n$ if $0 < \liminf a_n/b_n \leq \limsup_{n\to\infty} b_n/a_n < \infty$.

Figure 1: Error Convergence Rates n^{β} of Refined Approximation



Lemma 1 *Let* $k^* \asymp n^{\alpha^*}$ *with*

$$\alpha^* = \begin{cases} \max(\min\left(\frac{6\xi-1}{6\xi}, \frac{6\xi+2\delta-3}{6\xi+2\delta-1}\right), 0) & \text{for } 1/3 < \xi < 1/2 \\ \min\left(\frac{2\delta}{1+2(\delta+\xi)}, \frac{1}{1+\xi}\right) & \text{for } 1/2 \le \xi < 1. \end{cases}$$

Then $\min_{k \ge 1} R(k, n, \xi, \delta) \asymp R(k^*, n, \xi, \delta) \asymp n^{\beta^*}$ *with*

$$\beta^* = \begin{cases} -\delta & \text{for } \delta \le 3(1/2 - \xi) \\ -\frac{3+2\delta-6\xi}{12\xi+4\delta-2} & \text{for } 3(1/2-\xi) < \delta \le 1/2 + 3\xi \\ -\frac{1}{6\xi} & \text{for } 1/2 + 3\xi < \delta \end{cases}$$

for $1/3 < \xi < 1/2$*, and* $\beta^* = -\xi\alpha^*$ *for* $1/2 \le \xi < 1$.

Remarks.

1. For $1/3 < \xi < 1/2$, Hall (1979) shows that under Condition 1, the error in the usual normal approximation to the distribution of S_n satisfies $\sup_s |\mathbb{P}(n^{-1/2}S_n \le s) - \mathbb{P}(\sigma_0 Z \le s)| \asymp n^{1-1/(2\xi)}$, so convergence is very slow for ξ close to 1/2. For $\xi = 1/2$, Theorems 3 and 4 in Hall (1980) imply that under Condition 1, $(n \log n)^{-1/2} S_n$ converges to a normal distribution at a logarithmic rate. For any $\delta > 0$, the new approximation with optimal choice of k^* yields a better rate n^{β^*} for ξ sufficiently close to 1/2, and for sufficiently large δ, the rate is at least as fast as $n^{-1/3}$ for all $1/3 < \xi \le 1/2$. Thus, if the tail of F is sufficiently close to being Pareto in the sense of Condition 1, then the new approximations can provide dramatic improvements over the normal approximation. Even keeping k fixed improves over

the benchmark rate $n^{1-1/(2\xi)}$ as long as $\delta > 1/(2\xi) - 1$ for $1/3 < \xi < 1/2$. At the same time, if $\delta < 1/2$, then β^* is larger than $1 - 1/(2\xi)$ for some ξ sufficiently close to $1/3$, so the new approximation is potentially worse than the usual normal approximation (or, equivalently, the optimal choice of k then is $k^* = 0$).

For $1/2 < \xi < 1$ and under Condition 1, $\sup_s |\mathbb{P}(n^{-\xi} S_n \leq s) - \mathbb{P}(\omega \sum_{i=1}^{\infty} \Gamma_i^{-\xi} \leq s)| = O(n^{1-2\xi} + n^{-\delta})$ by Theorem 1 of Hall (1981), and his Theorem 2 shows this rate to be sharp under a suitably strengthened version of Condition 1. More specifically, for F mean-centered Pareto, the rate is exactly $n^{1-2\xi}$ (cf. Christoph and Wolf (1992), Example 4.25), which, for any $\delta > 0$, is slower than n^{β^*} for ξ sufficiently close to $1/2$.

Figure 1 plots some of these rates.

2. An alternative approximation is obtained by replacing the term in the positive part function in parts (a) and (b) of Theorem 1 by $\sigma^2(\omega(n/\Gamma_k)^\xi)$, with an approximation error that is still bounded by $C \cdot R(k, n, \xi, \delta)$. Substitution of the term $\sigma_0^2 - \frac{\omega^2}{1-2\xi}(\Gamma_k/n)^{1-2\xi}$ in part (a) of Theorem 1 by $\sigma_0^2 - \frac{\omega^2}{1-2\xi}(k/n)^{1-2\xi}$ (or dropping the integral in part (b) for $1/3 < \xi < 1/2$) induces an additional error of order $(k/n)^{1-2\xi} k^{-1/2}$. In general, this worsens the bound, although even with this further approximation, the rate can still be better than the baseline rate of $n^{1-1/(2\xi)}$. For $1/2 < \xi < 1$, dropping the integral in part (b) induces an additional error of order $k^{-\xi}$, so this simpler approximation still has an error no larger than $C \cdot R(k, n, \xi, \delta)$.

3. Consider the case where both tails of F are approximately Pareto, that is Condition 1 holds for $\xi = \xi_R$ and $\delta = \delta_R$, and for some $x_L, \omega_L, \delta_L, L_L > 0$, for all $x < -x_L$, $f(x) = (\omega_L \xi_L)^{-1}(-x/\omega_L)^{-1/\xi_L - 1}(1 + h_L(-x))$ with $|h_L(x)| \leq L_L x^{-\delta_L/\xi_L}$ for all $x > x_L$. Proceeding as in the introduction then suggests

$$S_n \overset{a}{\sim} n^{1/2}\sigma(\omega_L(n/\Upsilon_{k_L})^{\xi_L}, \omega_R(n/\Gamma_{k_R})^{\xi_R})Z + n^{\xi_L}\frac{\omega_L}{1-\xi_L}\Upsilon_{k_L}^{1-\xi_L} - n^{\xi_R}\frac{\omega_R}{1-\xi_R}\Gamma_{k_R}^{1-\xi_R} + n^{\xi}\omega_R\sum_{i=1}^{k_R}\Gamma_i^{-\xi_R} - n^{\xi_L}\omega_L\sum_{i=1}^{k_L}\Upsilon_i^{-\xi_L}$$

with $(\Upsilon_i)_{i=1}^{\infty}$ an independent copy of $(\Gamma_i)_{i=1}^{\infty}$ and $\sigma^2(x, y)$ the variance of X_1 conditional on $-x \leq X_1 \leq y$. If $1/3 < \xi_L, \xi_R < 1$, then arguments analogous to the proof of Theorem 1 show that the error of this approximation is bounded by an expression of the form $C \cdot R(k_R, n, \xi_R, \delta_R) + C \cdot R(k_L, n, \xi_L, \delta_L)$, and the same form is obtained by replacing $\sigma^2(x, y)$ with $\sigma^2(\omega_L x, \omega_R y) = (\sigma^2(\omega_L v_n, \omega_R u_n) + (\omega_L^2/\xi_L)\int_{v_n}^x t^{1-1/\xi_L} dt + (\omega_R^2/\xi_R)\int_{u_n}^y t^{1-1/\xi_R} dt)_+$ for $v_n = (n/k_L)^{\xi_L}$ and $u_n = (n/k_R)^{\xi_R}$ (and the integrals may be dropped for $1/2 < \xi < 1$, see the preceding remark). If $\bar{\xi} = \max(\xi_L, \xi_R) > 1/2$ and $\xi_L \neq \xi_R$, then the first order approximation to the distribution of $n^{-\bar{\xi}} S_n$ is a one-sided stable law that does not depend on the smaller tail index. In contrast, the approximation above reflects the impact of both heavy tails, and in general, ignoring the relatively lighter tail leads to a worse bound.

4. Suppose the right tail of F is well approximated by a shifted Pareto distribution, that is for some $\kappa \in \mathbb{R}$ and $x_1, \delta_1, L_1 > 0$, $dF(x)/dx = f(x) = (\omega\xi)^{-1}((x-\kappa)/\omega)^{-1/\xi-1}(1 + h(x-\kappa))$ for all $x > x_1 + \kappa$ with $|h(y)| \leq L_1 y^{-\delta_1/\xi}$ uniformly in $y \geq x_1$. This implies that F satisfies Condition 1, but only for $\delta = \min(\xi, \delta_1)$. Let $F_0(x) = F(x+\kappa)$ and $\mu_0(x) = -\int_x^\infty y dF_0(y)/F_0(x)$. Then $\mu(x+\kappa) = -\int_{x+\kappa}^\infty y dF(y)/F(x+\kappa) = [\mu_0(x) - \kappa(1 - F_0(x))]/F_0(x)$. Thus, proceeding as for (6) yields $(X_{n-i+1:n})_{i=1}^k \overset{a}{\sim} (\kappa + n^\xi \omega \Gamma_i^{-\xi})_{i=1}^k$ and

$$S_n \overset{a}{\sim} \kappa(k - \Gamma_k) - n^\xi \frac{\omega}{1-\xi} \Gamma_k^{1-\xi} + n^{1/2}\sigma(\omega(n/\Gamma_k)^\xi + \kappa)Z + n^\xi \omega \sum_{i=1}^k \Gamma_i^{-\xi}. \tag{9}$$

Straightforward modifications of the proof of Theorem 1 show that the approximation error in (9) is bounded by $C \cdot R(k, n, \xi, \delta_1)$, and this form for the bound also applies if $\sigma^2(\omega x + \kappa)$ is further approximated by $\sigma^2(\omega x + \kappa) \approx (\sigma^2(\omega u_n) + (\omega^2/\xi)\int_{u_n}^x y^{1-1/\xi} dy)_+$ for $u_n = (n/k)^\xi$. So, for instance, if F is mean-centered Pareto with $1/3 < \xi < 1$, then δ_1 may be chosen arbitrarily large, and the approximation (9) with $k = k^*$ of Lemma 1 yields a substantially better bound on the convergence rate compared to the original approximation (7) with a bound of the form $C \cdot R(k, n, \xi, \xi)$. The cost of this further refinement, however, is the introduction of a tail location parameter κ in addition to the tail scale and tail shape parameters (ω, ξ).

3 Proofs

Let $X_n^e = (X_{n-k+1:n}, X_{n-k:n}, \ldots, X_{n:n})$. The proof of Theorem 1 relies heavily on Corollary 5.5.5 of Reiss (1989) (also see Theorem 2.2.4 of Falk, Hüsler, and Reiss (2004)), which implies that under Condition 1,

$$\sup_{B^k} |\mathbb{P}(n^{-\xi}\omega^{-1} X_n^e \in B^k) - \mathbb{P}((\Gamma_k^{-\xi}, \Gamma_{k-1}^{-\xi}, \ldots, \Gamma_1^{-\xi}) \in B^k)| \leq C((k/n)^\delta k^{1/2} + k/n) \tag{10}$$

where the supremum is over Borel sets B^k in $\mathbb{R}^k$.

Without loss of generality, assume $x_0 > e$, $1 - (x_0/\omega)^{-1/\xi} > 0$ and $\sigma_0^2 - \omega^{1/\xi} x_0^{1-2\xi}/(1-2\xi) > 0$. We first prove two elementary lemmas. Let L denote a generic positive constant that does not depend on x or y, not necessarily the same in each instant it is used.

Lemma 2 *Under Condition 1, for all $x, y \geq x_0$,*

(a) for $1/3 < \xi < 1$, $|\mu(x)| \leq Lx^{1-1/\xi}$ and $\left|\mu(x) + \frac{\omega^{1/\xi}}{1-\xi}\frac{x^{1-1/\xi}}{1-(x/\omega)^{-1/\xi}}\right| \leq Lx^{1-(1+\delta)/\xi}$

(b) for $1/3 < \xi < 1/2$, $|\sigma^2(x) - \sigma_0^2 + \frac{\omega^{1/\xi}}{1-2\xi}x^{2-1/\xi}| \leq L(x^{2-(1+\delta)/\xi} + x^{2-2/\xi} + x^{-1/\xi})$

(c) for $1/2 < \xi < 1$, $L^{-1}x^{2-1/\xi} \leq \sigma^2(x) \leq Lx^{2-1/\xi}$ and $|\sigma^2(x) - \sigma^2(y) + (\omega^{1/\xi}/\xi)\int_x^y u^{1-1/\xi}du| \leq L(|\int_x^y u^{1-(1+\delta)/\xi}du| + (x^{2-1/\xi} + y^{2-1/\xi})(x^{-1/\xi} + y^{-1/\xi}))$

(d) for $\xi = 1/2$, $L^{-1}\log x \leq \sigma^2(x) \leq L\log x$, and $|\sigma^2(x) - \sigma^2(y) + (\omega^{1/\xi}/\xi)\int_x^y u^{1-1/\xi}du| \leq L(|\int_x^y u^{1-(1+\delta)/\xi}du| + x^{-1/\xi}\log y + y^{-1/\xi}\log x)$

(e) for $1/3 < \xi < 1$, $\int |y|^3 d\tilde{F}_x(y) \leq Lx^{3-1/\xi}$.

Proof. (a) Follows from $\mu(x) = -\int_x^\infty udF(u)/F(x)$ and, under Condition 1, $|\int_x^\infty udF(u) - \frac{\omega^{1/\xi}}{1-\xi}x^{1-1/\xi}| \leq \int_x^\infty u|h(u)|du \leq Lx^{1-(1+\delta)/\xi}$ and $|F(x) - 1 + (x/\omega)^{-1/\xi}| \leq \int_x^\infty |h(u)|du \leq Lx^{-(1+\delta)/\xi}$.

(b),(c),(d) Since $\sigma^2(x) = -\mu(x)^2 + (\sigma_0^2 - \int_x^\infty u^2dF(u))/F(x)$ for $\xi < 1/2$ and $\sigma^2(x) = -\mu(x)^2 + \int_{-\infty}^x u^2dF(u)/F(x)$ for $\xi \geq 1/2$, the results follow from $|1 - F(x)| \leq Lx^{-1/\xi}$, $|\int_x^y u^2dF(u) - (\omega^{1/\xi}/\xi)\int_x^y u^{1-1/\xi}du| \leq L\int_x^y u^{1-(1+\delta)/\xi}du$ via Condition 1 and the result in part (a).

(e) Follows from $\int |u|^3 d\tilde{F}_x(u) = \int_{-\infty}^x |u - \mu(x)|^3 dF(u)/F(x) \leq L|\mu(x)|^3 + L\int_{-\infty}^x |u|^3 dF(u)$ by the c_r inequality and Condition 1. ■

Lemma 3 *Under Condition 1*

(a) with $\tilde{T}_n = \max(T_n, x_0)$, $\mathbb{E}[\tilde{T}_n^\alpha] \leq C(n/k)^{\alpha\xi}$ for all $0 \leq \alpha < 1/\xi$

(b) with $\tilde{\tau}_n = \max(\omega n^\xi \Gamma_k^{-\xi}, x_0)$, $\mathbb{E}[\tilde{\tau}_n^{-\alpha}] \leq C(n/k)^{-\alpha\xi}$ for all $\alpha \geq 0$.

Proof. (a) Let $Y_n = (k/n)^\xi \tilde{T}_n$, so that we need to show that $\mathbb{E}[Y_n^\alpha]$ is uniformly bounded or, equivalently, that $\mathbb{P}(Y_n \geq y)y^{\alpha-1}$ is uniformly integrable. We have, for $y > x_0$

$$
\begin{aligned}
\mathbb{P}(Y_n \geq x) &= \mathbb{P}(T_n \geq (n/k)^\xi y) \\
&= \mathbb{P}(1 - F(T_n) \leq 1 - F((n/k)^\xi y)) \\
&\leq \mathbb{P}(U_{k:n} \leq \bar{L}y^{-1/\xi}k/n)
\end{aligned}
$$

where $U_{k:n}$ is the kth order statistic of n i.i.d. uniform $[0,1]$ variables, and $\bar{L}$ is such that $1 - F(x) \leq \bar{L}x^{-1/\xi}$ for all $x \geq x_0$. By Lemma 3.1.2 of Reiss (1989), for all $u > 0$, $\mathbb{P}(U_{k:n} \leq \frac{k}{n+1}u) \leq L(eu)^k$. Thus, $\mathbb{P}(Y_n \geq y) \leq L(\bar{L}y^{-1/\xi}e)^k \leq Ly^{-1/\xi}$, where the last inequality holds for all $y \geq (\bar{L}e)^\xi$, and the result follows.

(b) Clearly, $\mathbb{E}[\tilde{\tau}_n^{-\alpha}] \leq (n/k)^{-\alpha\xi}\mathbb{E}[(\Gamma_k/k)^{\alpha\xi}]$. For $0 \leq \alpha\xi \leq 1$, $\mathbb{E}[(\Gamma_k/k)^{\alpha\xi}] \leq \mathbb{E}[\Gamma_k/k]^{1/(\alpha\xi)} = 1$ while for $\alpha\xi > 1$, $\mathbb{E}[(\Gamma_k/k)^{\alpha\xi}] = \mathbb{E}[(k^{-1}\sum_{i=1}^k E_i)^{\alpha\xi}] \leq \mathbb{E}[k^{-1}\sum_{i=1}^k E_i^{\alpha\xi}] \leq C$ by two applications of Jensen's inequality. ■

Proof of Theorem 1

We can assume $k \leq n^{\frac{2\delta}{1+2\delta}}$ in the following, since otherwise, there is nothing to prove. Let $\tilde{T}_n = \max(T_n, x_0)$. Lemma 3.1.1 in Reiss (1989) implies that under Condition 1, $\mathbb{P}(\tilde{T}_n \neq T_n) \leq Ck/n$. Write $H_n(s) = \mathbb{P}(n^{-\gamma}S_n \leq s)$.

Assume first $1/3 < \xi \leq 1/2$. We have

$$H_n(s) = \mathbb{E}\left[\mathbb{P}\left(\sum_{i=1}^{n-k}\frac{X_{i:n}-\mu(T_n)}{(n-k)^{1/2}\sigma(T_n)} \leq \frac{s/a_n - \sum_{i=k+1}^n X_{i:n} - (n-k)\mu(T_n)}{(n-k)^{1/2}\sigma(T_n)}\middle| X_n^e\right)\right].$$

Note that conditional on X_n^e, the distribution of $\sum_{i=1}^{n-k} X_{i:n}$ is the same as that of the sum of i.i.d. draws from the truncated distribution $\tilde{F}_{T_n}$ with mean $\mu(T_n)$ and variance $\sigma(T_n)$. The Berry-Esseen bound hence implies

$$\sup_z\left|\mathbb{E}\left[\mathbf{1}[\sum_{i=1}^{n-k}\frac{X_{i:n}-\mu(T_n)}{(n-k)^{1/2}\sigma(T_n)} \leq z]|X_n^e\right] - \Phi(z)\right| \leq C(n-k)^{-1/2}\frac{\int|x|^3 d\tilde{F}_{T_n}(x)}{\sigma^3(T_n)}$$

where $\Phi(z) = P(Z \leq z)$. Replacing T_n by $\tilde{T}_n$, by Lemma 2 (e), $\int|x|^3 d\tilde{F}_{\tilde{T}_n}(x) \leq C(\tilde{T}_n)^{3-1/\xi}$ and $\sigma^3(\tilde{T}_n) \geq \sigma^3(x_0)$ a.s. From Lemma 3 (a), $\mathbb{E}[\tilde{T}_n^{3-1/\xi}] \leq C(n/k)^{3\xi-1}$, so that

$$\sup_s\left|H_n(s) - \mathbb{E}\Phi\left(\frac{s/a_n - \sum_{i=k+1}^n X_{i:n} - (n-k)\mu(\tilde{T}_n)}{(n-k)^{1/2}\sigma(\tilde{T}_n)}\right)\right| \leq C(n^{-1/2}(n/k)^{3\xi-1} + k/n).$$

From (10), with $\tau_n = \omega(n/\Gamma_k)^\xi$, $\sup_s|H_n(s) - H_n^1(s)| \leq C(n^{-1/2}(n/k)^{3\xi-1} + (k/n)^\delta k^{1/2} + k/n)$, where

$$H_n^1(s) = \mathbb{E}\Phi\left(\frac{s/a_n - \omega n^\xi\sum_{i=1}^k\Gamma_i^{-\xi} - (n-k)\mu(\tau_n)}{(n-k)^{1/2}\sigma(\tau_n)}\right).$$

Let $\tilde{\tau}_n = \max(\tau_n, x_0)$ and note that by (10), $P(\tilde{\tau}_n \neq \tau_n) \leq \mathbb{P}(\tilde{T}_n \neq T_n) + C((k/n)^\delta k^{1/2} + k/n)$.

Now focus on the claim in part (a). By Lemma 2 (a) and (b), $|\mu(\tilde{\tau}_n) + \frac{\omega^{1/\xi}}{1-\xi}\frac{\tilde{\tau}_n^{1-1/\xi}}{1-(\tilde{\tau}_n/\omega)^{-1/\xi}}| \leq C\tilde{\tau}_n^{1-(1+\delta)/\xi}$ and $|\sigma^2(\tilde{\tau}_n) - \sigma_0^2 + \frac{\omega^{1/\xi}}{1-2\xi}\tilde{\tau}_n^{2-1/\xi}| \leq C\max(\tilde{\tau}_n^{2-(1+\delta)/\xi}, \tilde{\tau}_n^{2-2/\xi}, \tilde{\tau}_n^{-1/\xi})$ a.s. Thus, exploiting that $\phi(z) = d\Phi(z)/dz$ and $|z|\phi(z)$ are uniformly bounded, and $0 < \sigma^2(x_0) \leq$

$\sigma^2(\tilde{\tau}_n) \leq \sigma_0^2$ a.s., exact first order Taylor expansions and Lemma 3 (b) yield

$$\sup_s \left| H_n^1(s) - \mathbb{E}\Phi\left(\frac{s - n^{\xi-1/2}\omega\sum_{i=1}^k \Gamma_i^{-\xi} - n^{\xi-1/2}\frac{\omega}{1-\xi}\Gamma_k^{1-\xi}\Psi_n}{\left(\sigma_0^2 - \frac{\omega^2}{1-2\xi}(\Gamma_k/n)^{1-2\xi}\right)^{1/2}}\right)\right|$$
$$\leq C((k/n)^\delta k^{1/2} + k/n + n^{1/2}(n/k)^{\xi-1-\delta} + (n/k)^{2\xi-(1+\delta)} + (n/k)^{2\xi-2})$$

where $\Psi_n = 1 + \frac{\Gamma_k/n - k/n}{1-\Gamma_k/n}$. Let $\tilde{\Gamma}_k = n(\tilde{\tau}_n/\omega)^{-1/\xi}$, so that $\mathbb{P}(\tilde{\Gamma}_k \neq \Gamma_k) = \mathbb{P}(\tilde{\tau}_n \neq \tau_n)$, and we can replace any Γ_k by $\tilde{\Gamma}_k$ in the last expression without changing the form of the right hand side. Note that $1 - \tilde{\Gamma}_k/n \geq 1 - (x_0/\omega)^{-1/\xi} > 0$ and $\sigma_0^2 - \frac{\omega^2}{1-2\xi}(\tilde{\Gamma}_k/n)^{1-2\xi} \geq \sigma^2(x_0)$ a.s. Thus, by another exact Taylor expansion and $\mathbb{E}[\Gamma_k^{1-\xi}|\Gamma_k/n - k/n|]^2 \leq \mathbb{E}[\Gamma_k^{2-2\xi}]\mathbb{E}[(\Gamma_k/n - k/n)^2] \leq Ck^{3-2\xi}/n^2$, we can replace Ψ_n by 1 at the cost of another error term of the form $C(n/k)^{-3/2+\xi}$. The result in part (a) now follows after eliminating dominated terms, and the proof of part (b) for $1/3 < \xi < 1/2$ follows from the same steps.

So consider $\xi = 1/2$. Let A_n be the event $(2k)^{-\xi} \leq \Gamma_k^{-\xi} \leq (k/2)^{-\xi}$. By Chebychev's inequality, $\mathbb{P}(A_n) = \mathbb{P}(1/2 \leq k^{-1}\sum_{i=1}^k E_i \leq 2) \leq C/k$. Conditional on A_n, and recalling that $k \leq n^{\frac{2\delta}{1+2\delta}}$, $C^{-1} \leq \sigma^2(\tilde{\tau}_n)/\log(n) \leq C$, $|\sigma^2(\tilde{\tau}_n) - \sigma^2(\omega u_n) - \frac{\omega^{1/\xi}}{\xi}\int_{\omega u_n}^{\tau_n} y^{1-1/\xi}dy| \leq C((n/k)^{2\xi-1-\delta} + (k/n)\log(n))$ and $|\mu(\tilde{\tau}_n) + \frac{\omega^{1/\xi}}{1-\xi}\frac{\tilde{\tau}_n^{1-1/\xi}}{1-(\tilde{\tau}_n/\omega)^{-1/\xi}}| \leq C(n/k)^{\xi-1-\delta}$ a.s. by Lemma 2 (a) and (d). Exact first order Taylor expansions of $H_n^1(s)$ thus yield

$$\sup_s \left| H_n^1(s) - \mathbb{E}\Phi\left(\frac{s(\log n)^{1/2} - \omega n^{1/2}\sum_{i=1}^k \Gamma_i^{-1/2} - n^{1/2}\frac{\omega}{1-\xi}\Gamma_k^{1/2}\Psi_n}{\left(\sigma^2(\omega u_n) + 2\omega^2\int_{u_n}^{(n/\Gamma_k)^{1/2}} y^{-1}dy\right)^{1/2}}\right)\right|$$
$$\leq C(k^{-1/2} + (k/n)^\delta k^{1/2} + k/n + n^{1/2}(n/k)^{-1/2-\delta} + (k/n)^{-\delta})$$

and replacing Ψ_n by unity induces an additional error term of the form $C(n/k)^{-1}$ by the same arguments as employed above (and recalling that $\mathbb{P}(A_n) \leq C/k$).

We are left to prove the claim for $1/2 < \xi < 1$. Note that the distribution of $\sum_{i=1}^{n-k} X_{i:n}$ conditional on X_n^e only depends on X_n^e through T_n. Let $\Phi_{n,t}$ be the conditional distribution function of $\sum_{i=1}^{n-k}\frac{X_{i:n} - \mu(T_n)}{(n-k)^{1/2}\sigma(T_n)}$ given $T_n = t$. For future reference, note that by Theorem 1.1 in Goldstein (2010), $||\Phi_{n,t} - \Phi||_1 = \int|\Phi(z) - \Phi_{n,t}(z)|dz \leq (n-k)^{-1/2}\int|y|^3 dF_t(y)/\sigma(t)^3$, so that by Lemma 2 (c) and (e), $||\Phi_{n,t} - \Phi||_1 \leq Cn^{-1/2}t^{1/(2\xi)}$ for $t \geq x_0$. We have

$$H_n(s) = \mathbb{E}\Phi_{n,T_n}\left(\frac{n^\xi s - \sum_{i=k+1}^n X_{i:n} - (n-k)\mu(T_n)}{(n-k)^{1/2}\sigma(T_n)}\right)$$

so that by (10), $\sup_s |H_n(s) - H_n^2(s)| \le C((k/n)^\delta k^{1/2} + k/n)$, where

$$H_n^2(s) = \mathbb{E}\Phi_{n,\tau_n}\left(\frac{n^\xi s - \omega n^\xi \sum_{i=1}^k \Gamma_i^{-\xi} - (n-k)\mu(\tau_n)}{(n-k)^{1/2}\sigma(\tau_n)}\right).$$

Let U be a uniform random variable on the unit interval, independent of $(\Gamma_i)_{i=1}^\infty$, and let $\Phi_{n,t}^{-1}$ be the quantile function of $\Phi_{n,t}$. Then

$$H_n^2(s) = \mathbb{P}\left(n^{-\xi}(n-k)^{1/2}\sigma(\tau_n)\Phi_{n,k,\tau_n}^{-1}(U) + n^{-\xi}(n-k)\mu(\tau_n) + \omega\sum_{i=1}^k \Gamma_i^{-\xi} \le s\right).$$

Since $\Gamma_1/\Gamma_2, \Gamma_2/\Gamma_3, \ldots, \Gamma_{k-1}/\Gamma_k, \Gamma_k$ are independent (cf. Corollary 1.6.11 of Reiss (1989)), the distribution of $(\Gamma_1/\Gamma_2)^{-\xi}$ conditional on $\Gamma_2, \Gamma_3, \ldots, \Gamma_k$ is the same as that conditional on Γ_2, which by a direct calculation is found to be Pareto with parameter $1/\xi$. Thus, with $G(z) = \mathbf{1}[z > 1](1 - z^{-1/\xi})$,

$$H_n^2(s) = \mathbb{E}G\left(\left[s - n^{-\xi}(n-k)^{1/2}\sigma(\tau_n)\Phi_{n,\tau_n}^{-1}(U) - n^{-\xi}(n-k)\mu(\tau_n) - \omega\sum_{i=2}^k \Gamma_i^{-\xi}\right]/(\omega\Gamma_2^{-\xi})\right).$$

Note that for arbitrary $a \ge 0$ and $y \in \mathbb{R}$, with $g(z) = dG(z)/dz$

$$\begin{aligned}
|\mathbb{E}G(y + a\Phi_{n,t}^{-1}(U)) - \mathbb{E}G(y + aZ)| &= |\int G(y + az)d(\Phi_{n,t}(z) - \Phi(z))| \\
&= a|\int (\Phi(z) - \Phi_{n,t}(z))g(y + az)dz| \\
&\le a\sup_y |g(y)| \cdot ||\Phi_{n,t} - \Phi||_1
\end{aligned}$$

where the second equality stems from Riemann-Stieltjes integration by parts. Conditional on the event A_n as defined above, $||\Phi_{n,\tilde{\tau}_n} - \Phi||_1 \le Ck^{-1/2}$, $C^{-1}(n/k)^{2\xi-1} \le \sigma^2(\tilde{\tau}_n) \le C(n/k)^{2\xi-1}$, $|\sigma^2(\tilde{\tau}_n) - \sigma^2(\omega u_n) - \frac{\omega^{1/\xi}}{\xi}\int_{\omega u_n}^{\tilde{\tau}_n} y^{1-1/\xi}dy| \le C((n/k)^{2\xi-1-\delta} + (n/k)^{2\xi-2})$ and $|\mu(\tilde{\tau}_n) + \frac{\omega^{1/\xi}}{1-\xi}\frac{\tilde{\tau}_n^{1-1/\xi}}{1-(\tilde{\tau}_n/\omega)^{-1/\xi}}| \le C(n/k)^{\xi-1-\delta}$ a.s. by Lemma 2 (a) and (c). Thus, by exact first order Taylor expansions and exploiting that $g(z)$ is uniformly bounded and $\mathbb{E}[|Z|]$, $\mathbb{E}[\Gamma_2^\xi] < C$,

$$\begin{aligned}
\sup_s |H_n^2(s) - H_n^3(s)| \le C(&k^{-1} + (k/n)^\delta k^{1/2} + k/n \\
&+ k^{-\xi} + n^{1-\xi}(n/k)^{\xi-1-\delta} + n^{1/2-\xi}((n/k)^{\xi-1/2-\delta} + (n/k)^{\xi-3/2}))
\end{aligned}$$

where

$$H_n^3(s) = \mathbb{E}G\left(\left[s - n^{1/2-\xi}\left(\sigma^2(\omega u_n) + \frac{\omega^{1/\xi}}{\xi}\int_{\omega u_n}^{\tau_n} y^{1-1/\xi}dy\right)^{1/2} Z \right.\right.$$
$$\left.\left. + \frac{\omega}{1-\xi}\Gamma_k^{1-\xi}\Psi_n - \omega\sum_{i=2}^{k}\Gamma_i^{-\xi}\right] / (\omega\Gamma_2^{-\xi})\right).$$

As before, we can replace Ψ_n by unity at the cost of another error term of the form $Ck^{3/2-\xi}/n$, and the result follows after eliminating dominating terms.

References

Christoph, G., and W. Wolf (1992): *Convergence theorems with a stable limit law*, Mathematical Research. Akademie Verlag, Berlin.

Csörgö, S., E. Haeusler, and D. M. Mason (1988): "A probabilistic approach to the asymptotic distribution of sums of independent, identically distributed random variables," *Advances in Applied Mathematics*, 9(3), 259–333.

de Haan, L., and S. Resnick (1996): "Second-order regular variation and rates of convergence in extreme-value theory," *The Annals of Probability*, 24(1), 97–124.

de Haan, L., and U. Stadtmüller (1996): "Generalized regular variation of second order," *Journal of the Australian Mathematical Society*, 61(3), 381–395.

Falk, M., J. Hüsler, and R. Reiss (2004): *Laws of Small Numbers: Extremes and Rare Events*. Birkhäuser, Basel.

Falk, M., and F. Marohn (1993): "Von Mises conditions revisited," *The Annals of Probability*, 21(3), 1310–1328.

Goldstein, L. (2010): "Bounds on the constant in the mean central limit theorem," *The Annals of Probability*, 38(4), 1672–1689.

Hall, P. (1979): "On the rate of convergence in the central limit theorem for distributions with regularly varying tails," *Probability Theory and Related Fields*, 49(1), 1–11.

——— (1980): "Characterizing the rate of convergence in the central limit theorem. I," *The Annals of Probability*, 8(6), 1037–1048.

——— (1981): "Two-sided bounds on the rate of convergence to a stable law," *Probability Theory and Related Fields*, 57(3), 349–364.

——— (1982): *Rates of convergence in the central limit theorem*. Pitman Publishing, Boston.

LePage, R., M. Woodroofe, and J. Zinn (1981): "Convergence to a stable distribution via order statistics," *The Annals of Probability*, 9(4), 624–632.

Reiss, R.-D. (1989): *Approximate distributions of order statistics: with applications to nonparametric statistics*. Springer Verlag, New York.