A Higher Order Expansion for the Joint Density of the Sum and the Maximum under the Gumbel Domain of Attraction *

Amy Grady
National Institute of Statistical Science

Richard L. Smith
Department of Statistics
University of North Carolina at Chapel Hill

April 16, 2003

Abstract

An asymptotic expansion is developed for the joint density of the sum and maximum of an *i.i.d.* sequence when the parent distribution is in the domain of extreme value attraction of the Gumbel distribution. Previous results by Chow and Teugels, extended by Anderson and Turkman, show that in this situation, the normalized sum and normalized maximum of the sample converge to independent normal and Gumbel distributions, but they have not characterized the rate of convergence. The present development proceeds via three technical propositions. The first extends previous results by Cohen and by Smith to derive the rate of convergence of the density of the sample maximum to a limiting Gumbel density. The second technical proposition is a conditional Edgeworth expansion for the sum given the maximum. The third concerns the expansion of conditional means and variances. By combining these propositions, a leading-term expansion is developed for the dependence between the sum and the maximum, and uniform convergence is proved over an expanding sequence of subsets of the plane. Simulations allow us to assess the practical applicability of the results. They show that for moderate sample sizes, the sum and the maximum are far from being independent, but that the leading-term asymptotic expansion is a substantial improvement over independence of the two random variables.

1 Introduction

This work extends the research into the relationship between the sum and maximum, specifically, the work of Chow and Teugels(1978) with the *iid* case – followed by Anderson and Turkman(1991) with the stationary case – who established that, under certain conditions on the domains of attraction, the sum and the maximum are asymptotically independent. They showed under appropriate conditions that the joint distribution function of the normalized sum and the normalized maximum converges to the product of their appropriate asymptotic distributions: a stable law for the sum

*Amy Grady is Research Fellow, National Institute of Statistical Sciences, 19 T.W. Alexander Drive, Research Triangle Park, NC 27739; agrady@niss.org. Richard L. Smith is Professor, Department of Statistics, University of North Carolina, Chapel Hill, NC 27599-3260, USA; email rls@email.unc.edu. The work of Amy Grady was supported in part by EPA Cooperative Agreement R-83042501-0. Richard L. Smith was supported in part by NSF grants DMS-9971980 and DMS-0084375. The paper is based on part of the PhD dissertation of Amy Grady at the Department of Statistics, University of North Carolina.

multiplied by one of the three extreme value distributions for the maximum. A chief theoretical question that remains is what is the rate of convergence to this asymptotic independence. The associated statistical methodology question is, for moderate sample size, how does one model the dependence structure between the sum and the maximum. Given the importance of both the rate of convergence and the statistical modelling questions, an important development in this area is a higher order *expansion* for the joint *density* of the sum and the maximum.

The primary goal of this paper is to establish a higher order term for this joint density when the maximum lies in the Gumbel domain of attraction. Here the focus is on the density, as opposed to the distribution function, for direct applications into likelihood functions. The other domains of attraction for the maximum, the Fréchet and the Weibull, are covered in Grady and Smith (2003). Statistical applications in climatology will be developed elsewhere.

1.1 Notation

Let $X_1, \ldots, X_n$ be an *iid* sequence of random variables with common distribution function F which has density f and characteristic function φ where the support of F lies on (x_l, x_o) where $-\infty \leq x_l$, $x_o \leq \infty$. We assume the existence of the third moment μ_3 from which follows the existence of the mean μ, variance σ^2, and third cumulant $\mathcal{K}_3$.

Define $S_n = \sum_{i=1}^n X_i$ with the normalized version as

$$S_n^* = \frac{S_n - n\mu}{\sqrt{n\sigma^2}}, \quad \text{where } \mu \text{ real and } \sigma > 0 \tag{1}$$

and $M_n = \max_{1 \leq i \leq n} X_i$ with the normalized version as

$$M_n^* = \frac{M_n - b_n}{a_n}, \quad \text{where } a_n > 0, b_n \text{ real.} \tag{2}$$

We define the distribution function of S_n^* as $F_{S_n^*}$ with density $f_{S_n^*}(w) = dF_{S_n^*}(w)/dw$. We also define the distribution function of M_n^* as $F_{M_n^*}(v) = F^n(a_n v + b_n)$ with density $f_{M_n^*}(v) = dF_{M_n^*}(v)/dv = nF^{n-1}(a_n v + b_n) f(a_n v + b_n) a_n$. To simplify let $u_n = a_n v + b_n$. Finally, we denote the joint density of S_n^* and M_n^* as $f_{S_n^*, M_n^*}(w, v)$.

1.2 Main Result

We use three sets of conditions – one set for deriving the expansion of the density of the maximum, another set for deriving the expansion of the conditional density of sum given the maximum, and one extra condition necessary for deriving the expansions of the appropriate moments. Under these conditions, defining

$$r_n = \begin{cases} \frac{b_n}{\sqrt{n\sigma^2}}, & x_o = \infty \\ \frac{x_o - \mu}{\sqrt{n\sigma^2}} & x_o < \infty \end{cases}$$

we have

$$|f_{S_n^*,M_n^*}(w,v) - f_{S_n^*}(w) f_{M_n^*}(v)\{1 - r_n(e^{-v} - 1)w\}| = o(r_n)$$

uniformly $\forall w$ and $\forall |v| \leq e_n$ where e_n is given explicitly in Proposition 1 and tends to ∞ as $n \to \infty$. In particular, the rate of convergence to asymptotic independence is of the order $O(\frac{b_n}{\sqrt{n}})$ for $x_o = \infty$ and $O(\frac{1}{\sqrt{n}})$ for $x_o < \infty$.

We derive two corollaries under the same set of conditions as in the main theorem. In the first of these corollaries, we substitute in the appropriate asymptotic densities for $f_{M_n^*}$ and $f_{S_n^*}$ – namely, the Gumbel density and the normal density, respectively. We conclude that the leading order expansion terms in this case not only include the term in the main result which models a first order approximation to the dependence between the sum and the maximum but also, depending on the underlying distribution, could incorporate the expansion terms associated with the density of the maximum and/or Edgeworth expansion terms associated with the density of the sum. A second corollary to the main theorem establishes a higher order expansion for the joint density for the penultimate approximation case. In this case, the resulting higher order terms are identical to those in the previous results.

Our approach in solving this higher order expansion for the joint density is to rewrite $f_{S_n^*,M_n^*}(w,v) = f_{S_n^*|M_n^*}(w|v) \times f_{M_n^*}(v)$ where $f_{S_n^*|M_n^*}(w|v)$ is the conditional density of S_n^* given M_n^*. Then we need to establish three key expansions. These expansions form three main propositions of this paper which in turn are central to establishing the main result. In the first of these proposition (Proposition 1) we derive the expansion for $f_{M_n^*}$. In the second (Proposition 3), we establish the expansion for a transformation of $f_{S_n^*|M_n^*}$. Finally in the third (Proposition 4), we derive the expansions for the appropriate *conditional* moments.

1.3 Proposition 1: The expansion of the density of the maximum

We assume that F lies in the Gumbel domain of attraction. We let Λ denote the Gumbel distribution function and Λ' denote its density where

$$\Lambda(x) = e^{-e^{-x}}, \ -\infty < x < \infty.$$

Following Smith (1987), we assume F has the representation

$$-\log\ F(x) = c(x) \exp\left\{-\int_{-\infty}^{x} \frac{dt}{\phi(t)}\right\} \quad \forall x < x_o \tag{3}$$

where ϕ is a positive twice differentiable function, $\phi'(x) \to 0$ and $c(x) \to 1$ as $x \to x_o$, and $x_o = \sup\ \{x : 1 - F(x) > 0\}$. This representation, originally due to Balkema and de Haan (1972), is an adaptation of a Karamata representation for $F \in \mathcal{D}(\Lambda)$.

The normalizing constants in (2) are defined by

$$b_n = \inf \{x : -\log \ F(x) > 1/n\} \tag{4}$$

and

$$a_n = \phi(b_n). \tag{5}$$

For this paper, we also assume that $c(x) \equiv 1$ in (3). This is equivalent to the twice differentiable domain of attraction of Pickands (1986) for the Gumbel case which is equivalent to the von Mises condition for the Gumbel case [see Leadbetter *et al.* (1983), Theorem 1.6.1] which are the classic sufficient conditions for the local domain problem. Most well-behaved distributions in the Gumbel domain of attraction which have differentiable densities also satisfy the von Mises condition so in this case the assumption $c(x) \equiv 1$ is justifiable.

Finally, we also assume conditions on the function ϕ in (3). This function ϕ describes how smooth the tail of the underlying distribution is. Conditions placed on this function and its derivatives play a central role not only in the local domain problem; but also in establishing convergence rates for the distribution and density limits and higher order terms.

Rates of convergence to the Gumbel distribution for sample maxima were established by Anderson (1971) under general conditions but without uniformity results, and in specific cases by Hall and Wellner (1979), Hall (1979,1980), and Cohen (1982b). Cohen (1982a) establishes a general result, with uniformity, for two classes of distribution that he labelled N and E. Smith (1987) used a variation on class N defined by (3) with specific restrictions on ϕ. Under conditions similar to Proposition 9.2 of Smith (1987) and one other technical condition, Proposition 1 shows that

$$\left| f_{M_n^*}(v) - \Lambda'(v)\{1 + (\frac{v^2}{2} - v - \frac{v^2 e^{-v}}{2})\phi'(b_n) \right| = o(\phi'(b_n)) \quad \forall v.$$

Although this higher order term does not contribute to the form the main result takes, this term does contribute to the main result's first corollary. In fact, we substitute $\Lambda'(v)\{1+(\frac{v^2}{2} - v - \frac{v^2 e^{-v}}{2})\}\phi'(b_n)$ for $f_{M_n^*}(v)$ into the main result to help establish its first corollary. More specifically, our complete result in Proposition 1 provides – uniformly over the entire range – not only an expansion with first and second order terms for the density of the maximum but also an explicit form of this rate of convergence.

1.4 Proposition 3: The expansion of the conditional density of the sum given the maximum

The derivation we establish for the expansion of the conditional density for the sum given the maximum is patterned after the unconditional Edgeworth expansion found in Feller (1971), Chapter XVI, Section 2, Theorem 1 which gives the uniform expansion of the density of S_n^*. We make use not only of this theorem but also its supporting lemmas, modifying them for the conditional case.

In the development of the higher order expansion for the joint density, it will also be necessary to bound the density of the sum multiplied by a polynomial in order to establish the uniformity of the results. Under the same conditions of Feller (1971), Chapter XVI, Section 2, Theorem 1, Petrov (1975) gives the needed result.

Theorem 17 of Petrov (1975) for "k=3": Let $\{X_n\}$ be a sequence of independent random variables having a common distribution with zero mean, non-zero variance, and $E|X|^3 < \infty$. Let the random variable $\frac{1}{\sqrt{n\sigma^2}} S_n$ have for some $n > N$ a bounded density $f_n(x)$. Then

$$(1 + |x|^3) \left\{ f_n(x) - \mathcal{N}'(x) - \frac{\mu_3}{6\sigma^3\sqrt{n}}(x^3 - 3x)\mathcal{N}'(x) \right\} = o(1/\sqrt{n})$$

uniformly in x.

In our derivation, we begin by conditioning S_n on $M_n = u_n$ where $u_n = a_n v + b_n$ with a_n and b_n defined in (5) and (4) and with v fixed. This conditional distribution may be written in the form

$$P[S_n \leq x | M_n = u_n] = P\left[(\sum_{i=1}^{n-1} X_i^* + u_n) \leq x | M_n = u_n \right]$$

where X_i^*s are *iid* conditional random variables with

$$P[X_i^* \leq x] = P[X \leq x | X \leq u_n].$$

Henceforth we write X^* for the random variable with the distribution $F(x)/F(u_n), \forall x \leq u_n$ and denote its density as f_{u_n} and its characteristic function as φ_{u_n}. Then we write $\overline{S_n} = X_i^* + u_n$ where $X_1^*, \ldots, X_{n-1}^*$ are *iid* with the same distribution as X^*. The dependence on a given sequence $\{u_n\}$ – actually, sequences $\{a_n\}$ and $\{b_n\}$ – is implicit in the notation. Thus the conditional distribution of S_n given $M_n = u_n$ is the same as the unconditional distribution of $\overline{S_n}$, and we shall use this equivalence in the following discussion.

We may write

$$E(\overline{S_n}) = (n-1)\mu(u_n) + u_n$$

where $\mu(u_n) = E(X^*) = E(X|X \leq u_n)$ and

$$Var(\overline{S_n}) = (n-1)\sigma^2(u_n)$$

where $\sigma^2(u_n) = Var(X^*) = Var(X|X \leq u_n)$. Let $\kappa_3(u_n)$ denote the third cumulant of X^*. Expansions for $\mu(u_n)$ and $\sigma^2(u_n)$ are developed in Proposition 4.

Let $\tilde{S_n} = \{\overline{S_n} - E(\overline{S_n})\}/\sqrt{var(\overline{S_n})}$ denote the normalized version of $\overline{S_n}$. Then

$$P[\tilde{S_n} \leq x] \quad = \quad P\left[\frac{\overline{S_n} - \{(n-1)\mu(u_n) + u_n\}}{\sqrt{(n-1)\sigma^2(u_n)}} \leq x \right]$$

$$
\begin{aligned}
&= P\left[\frac{S_n - \{(n-1)\mu(u_n) + u_n\}}{\sqrt{(n-1)\sigma^2(u_n)}} \le x | M_n = u_n\right] \\
&= P\left[\frac{S_n - \{(n-1)\mu(u_n) + u_n\}}{\sqrt{(n-1)\sigma^2(u_n)}} \le x | M_n^* = v\right] \\
&= P\left[\frac{S_n^*\sqrt{n\sigma^2} + n\mu - \{(n-1)\mu(u_n) + u_n\}}{\sqrt{(n-1)\sigma^2(u_n)}} \le x | M_n^* = v\right] \qquad (6)
\end{aligned}
$$

where the last step comes from substituting (1) into the formula. Thus (6) gives the form of the transformation necessary when changing the variable $S_n^*|M_n^*$ to $\tilde{S_n}$. The Jacobian of this transformation is $\frac{\sqrt{n\sigma^2}}{\sqrt{(n-1)\sigma^2(u_n)}}$. In establishing the main result we actually use $f_{S_n^*,M_n^*} = f_{S_n^*|M_n^*} f_{M_n^*} = f_{\tilde{S_n}} f_{M_n^*} \frac{\sqrt{n\sigma^2}}{\sqrt{(n-1)\sigma^2(u_n)}}$.

Specifically, Proposition 3 derives a higher order expansion for the density of the sum normalized with respect to the conditional moments under a set of conditions similar to those in Feller (1971), Chapter XVI, Section 2, Theorem 1. These conditions are modified to deal with the conditioning on the maximum as defined in (6). In Proposition 3 we show that the conditional Edgeworth expansion of $f_{\tilde{S_n}}$ is

$$
f_{\tilde{S_n}}(x) - \mathcal{N}'(x) - \frac{\mathcal{K}_3(u_n)}{6\sigma^3(u_n)\sqrt{n}}(x^3 - 3x)\mathcal{N}'(x) = o(\frac{1}{\sqrt{n}}) \text{ uniformly in } x \text{ and } v
$$

where the dependence on v is through the definition of u_n. These Edgeworth terms do not contribute to the form of the main result. They are important in the corollaries, particularly when $x_o < \infty$ where the leading order Edgeworth term and the leading order term in the main result are both $O(1/\sqrt{n})$.

1.5 Proposition 4: The expansions of the conditional moments

The last proposition establishes the expansion terms for the difference between the unconditional and conditional means, $\mu - \mu(u_n)$, and the unconditional and conditional variances, $\sigma^2 - \sigma^2(u_n)$. Thus Proposition 4 provides a correction term for these moments when incorporating the information on whether the values are at or below the value u_n. This correction term for the first moments is particularly important because it directly contributes to the higher order term obtained in the main result and hence to the term developed for statistical modelling.

The first step in deriving these correction terms involves a law of total probability which introduces the conditional moments: those at or below u_n $[\mu(u_n) and \sigma^2(u_n)]$ and those above u_n, an exceedance. Here, we let Y to be an exceedance of u_n, i.e. $Y = (X - u_n)|X > u_n$, with mean $m(u_n)$ and variance $s^2(u_n)$. In doing so we make use of the relationship between the maxima and exceedances – in particular, Pickands (1975) shows that the distribution of the maximum converges to an extreme value distribution under the same conditions that the distribution of an (associated) exceedance converges to a generalized Pareto distribution. In particular, each of the necessary expansions of the terms in the low of total probability rely on the Balkema and de Haan representation (3) and Cohen's (1987b) class N conditions on ϕ – (17) and (18) – where Cohen's conditions hold on the interval $|v| \leq e_n$. Thus using the conditions of Proposition 1 [a variant on Section 9.2 of Smith(1987)], we obtain the necessary expansions for those components of $\mu - \mu(u_n)$ and $\sigma - \sigma(u_n)$ on that interval.

Then we introduce an additional technical condition when $x_o = \infty$ so as to guarantee that $\frac{u_n}{b_n} \to 1$ uniformly on $|v| \leq e_n$. On this interval, we can replace b_n for u_n when $x_o = \infty$. Finally, after comparing the terms of the expansions we have obtained, Proposition 4 provides $\forall |v| \leq e_n$

$$\mu - \mu(u_n) \sim \begin{cases} \frac{b_n e^{-v}}{n} & x_o = \infty \\ \frac{(x_o - \mu)e^{-v}}{n} & x_o < \infty \end{cases} \tag{7}$$

and

$$\sigma^2 - \sigma^2(u_n) \sim \begin{cases} \frac{b_n^2 e^{-v}}{n} & x_o = \infty \\ \{(x_o - \mu)^2 - \sigma^2\}\frac{e^{-v}}{n} & x_o < \infty. \end{cases} \tag{8}$$

1.6 Heuristic development of main result

Here we outline a heuristic development of the main result – the higher order expansion of the joint density of the sum and the maximum – when $x_o = \infty$, in order to give an appreciation for where the individual components in the higher order term originate and how they interact.

We begin by writing

$$f_{S_n^*, M_n^*}(w, v) = f_{S_n^*|M_n^*}(w|v) f_{M_n^*}(v).$$

By using the transformation defined in (6),

$$f_{S_n^*, M_n^*}(w, v) = f_{\tilde{S_n}}(z) \sqrt{\frac{n\sigma^2}{(n-1)\sigma^2(u_n)}} f_{M_n^*}(v) \tag{9}$$

where

$$z = \frac{\sqrt{n\sigma^2}w + n\mu - (n-1)\mu(u_n) - u_n}{\sqrt{(n-1)\sigma^2(u_n)}}. \tag{10}$$

To establish the correct form of this expansion, we multiply and divide (9) by $f_{S_n^*}$,

$$f_{S_n^*, M_n^*}(w, v) = f_{S_n^*}(w) f_{M_n^*}(v) \frac{f_{\tilde{S_n}(z)}}{f_{S_n^*}(w)} \sqrt{\frac{n\sigma^2}{(n-1)\sigma^2(u_n)}}. \tag{11}$$

Focusing on the ratio in the third term, we can express both the numerator and denominator using their appropriate Edgeworth expansions:

$$f_{S_n^*}(w) = \mathcal{N}'(w) \left\{ 1 + \frac{\kappa_3}{6\sigma^3\sqrt{n}}(w^3 - 3w) \right\} + o(1/\sqrt{n})$$

uniformly $\forall w$ by the standard Edgeworth expansion theory – see Feller (1971), Chapter XVI, Section 2, Theorem 1 – and

$$f_{\tilde{S_n}}(z) = \mathcal{N}'(z)\left\{1 + \frac{\kappa_3(u_n)}{6\sigma^3(u_n)\sqrt{n}}(z^3 - 3z)\right\} + o(1/\sqrt{n})$$

uniformly $\forall z$ by the conditional Edgeworth expansion derived in this paper in Proposition 3.

To simplify the $\frac{\mathcal{N}'(z)}{\mathcal{N}'(w)}$ term in the ratio, we use a Taylor's expansion of $\mathcal{N}'(z)$ about $\mathcal{N}'(w)$:

$$\mathcal{N}'(z) \approx \mathcal{N}'(w) + (z - w)\mathcal{N}''(w).$$

Substituting in the identity $\mathcal{N}''(w) = -w\mathcal{N}'(w)$,

$$\frac{\mathcal{N}'(z)}{\mathcal{N}'(w)} \approx 1 - (z - w)w. \tag{12}$$

Substituting (12) and the ratio of the remain Edgeworth expansion terms back into (11) yields:

$$f_{S_n^*,M_n^*}(w,v) \approx f_{S_n^*}(w)f_{M_n^*}(v)\{1 - (z - w)w\} \times \{\frac{1 + \frac{\kappa_3(u_n)}{6\sigma^3(u_n)\sqrt{n}}(z^3 - 3z)}{1 + \frac{\kappa_3}{6\sigma^3\sqrt{n}}(w^3 - 3w)}\}\sqrt{\frac{n\sigma^2}{(n-1)\sigma^2(u_n)}}. \tag{13}$$

Using the definition of z in (10),

$$\begin{aligned} z - w &= \frac{n(\mu - \mu(u_n))}{\sqrt{(n-1)\sigma^2(u_n)}} + \left\{\sqrt{\frac{n\sigma^2}{(n-1)\sigma^2(u_n)}} - 1\right\}w \\ &\quad + \frac{\mu(u_n)}{\sqrt{(n-1)\sigma^2(u_n)}} - \frac{u_n}{\sqrt{(n-1)\sigma^2(u_n)}}. \end{aligned} \tag{14}$$

Note (8) implies

$$\sqrt{\frac{n\sigma^2}{(n-1)\sigma^2(u_n)}} - 1 \sim \frac{b_n^2 e^{-v}}{n\sigma^2}.$$

Then Proposition 4 leads to

$$\begin{aligned} z - w &\sim \frac{b_n e^{-v}}{\sqrt{n\sigma^2}} + \frac{b_n^2 e^{-v}}{n\sigma^2}w + \frac{\frac{b_n e^{-v}}{\sqrt{n}} - \mu}{\sqrt{n\sigma^2}} - \frac{(a_n v + b_n)}{\sqrt{n\sigma^2}} \\ &\approx \frac{b_n(e^{-v} - 1)}{\sqrt{n\sigma^2}}. \end{aligned} \tag{15}$$

In the above approximation, we use the result $\frac{a_n}{b_n} \to 0$ to go to the second step. This result – a version of (126) of Lemma 16 – derives from the Balkema and de Haan representation (3).

Thus the higher order term of the joint density is $\frac{b_n(e^{-v}-1)}{\sqrt{n\sigma^2}}w$ and we conclude by substituting (15) into (13)

$$f_{S_n^*,M_n^*}(w,v) \approx f_{S_n^*}(w)f_{M_n^*}(v)\ \{1 - \frac{b_n(e^{-v} - 1)}{\sqrt{n\sigma^2}}w\}.$$

which confirms our main result when $x_o = \infty$. Note that in (13) the ratio of Edgeworth expansion and the Jacobian term have second and third moment terms thus neither contribute to the higher order expansion of the joint density.

1.7 Outline of paper

To develop the main theorem this paper is divided into six more parts. Section 2 presents Proposition 1 and its proof for the expansion for the density of the maximum under the Gumbel domain of attraction. Section 3 contains Proposition 3 and its proof which derives the expansion for the conditional density of the sum given the maximum. Section 4 is Proposition 4 and its proof for the expansions for the difference between the unconditional and conditional mean and variances. Finally, Section 5 presents the main theorem for the expansion of the joint density, Theorem 5, and give its proof along with two corollaries and their proofs. Section 6 presents a simulation project. Concluding remarks are in Section 7.

2 Expansion of the Density of the Maximum

Proposition 1 derives a higher order expansion for the density of M_n^*, the maximum of an *iid* sequence, when the underlying distribution of the observations, F, lies in the domain of attraction of the Gumbel distribution, Λ. Appendix A contains the lemmas that support this proposition.

Proposition 1 *Suppose $F \in \mathcal{D}(\Lambda)$ such that the representation in (3) holds with the definitions of a_n and b_n from (5) and (4). Assume for a constant $K > 2$,*

$$a_n v + b_n \to x_o \quad \textit{uniformly in } |v| \leq e_n = -K \log\ |\phi'(b_n)| \textit{ as } n \to \infty. \tag{16}$$

$$\phi''(a_n v + b_n)/\phi''(b_n) \to 1 \quad \textit{uniformly in } |v| \leq e_n = -K \log\ |\phi'(b_n)| \textit{ as } n \to \infty. \tag{17}$$

$$\phi(b_n)\phi''(b_n) \log\ |\phi'(b_n)|/\phi'(b_n) \to 0 \quad \textit{as } n \to \infty. \tag{18}$$

Suppose also that for any $v^* < x_o$

$$\inf_{v \leq v^*} \frac{f'(v)F(v)}{f^2(v)} > -\infty. \tag{19}$$

Then for each $\delta > 0$ there exists an n^* and a function ϵ_{n^*} tending to 0 as $n \to \infty$ such that $\forall n > n^*$ and $\forall v$

$$\begin{aligned} &|f_{M_n^*}(v) - \Lambda'(v)\{1 + (\frac{v^2}{2} - v - \frac{v^2 e^{-v}}{2})\phi'(b_n) \\ &\quad + (\frac{v^2}{2} + \frac{v^3}{6}[e^{-v} - 1])(2\phi'(b_n)^2 - \phi(b_n)\phi''(b_n)) \\ &\quad\quad + (\frac{v^3}{2}[e^{-v} - 2] + \frac{1}{8}v^4[e^{-2v} - 3e^{-v} + 1])\phi'(b_n)^2\}| \\ &\quad\quad\quad < \epsilon_{n^*}[\phi'(b_n)^2 + |\phi(b_n)\phi''(b_n)| \min(1, |v|^{-\delta})]. \end{aligned} \tag{20}$$

Remark Condition (16) is part of the definition of Cohen's (1982a) class N, see equation (1.24) of Cohen (1982a). Conditions (17) and (18) are equations (1.59) and (1.60) in Theorem 9 of Cohen(1982a) which list sufficient conditions for his class N. In that paper, he lists in Table 1 many distributions which belong to class N, for example, the normal and lognormal. In other words, there exists a deep pool of distributions satisfying the conditions in Proposition 1.

Proof: Given F is continuous, there exists a b_n such that $-\log\ F(b_n) = 1/n$, so

$$F_{M_n^*}(v) = F^n(a_n v + b_n) = \exp\left\{-\frac{-\log\ F(a_n v + b_n)}{-\log\ F(b_n)}\right\}.$$

Using (3) with $c(v) = 1$,

$$\begin{aligned} F_{M_n^*}(v) &= \exp\left\{-\exp\left(-\int_{b_n}^{a_n v + b_n} \frac{dt}{\phi(t)}\right)\right\} \\ &= \exp\left\{-\exp\left(-\int_0^v \frac{\phi(b_n)}{\phi(a_n t + b_n)} dt\right)\right\} \end{aligned}$$

$\forall\ a_n v + b_n < x_o$ where a_n is defined in (5) and b_n is defined above.

Thus

$$\begin{aligned} f_{M_n^*}(v) &= \frac{dF_{M_n^*}(v)}{dv} \\ &= \exp\left\{-\exp\left(-\int_0^v \frac{\phi(b_n)}{\phi(a_n t + b_n)} dt\right)\right\} \times \exp\left\{-\int_0^v \frac{\phi(b_n)}{\phi(a_n t + b_n)} dt\right\} \times \frac{\phi(b_n)}{\phi(a_n v + b_n)} \end{aligned}$$

or, say,

$$f_{M_n^*}(v) = f_1(v) \times f_2(v) \times f_3(v) \quad \forall v < x_o.$$

First, we restrict our attention to the interval $|v| \leq e_n = -K \log\ |\phi'(b_n)|$.

The expansion of f_2 Since $f_2(v) = \exp\{-\int_0^v \frac{\phi(b_n)}{\phi(a_n t + b_n)} dt\}$, by a similar argument that led to equation (9.20) of Smith (1987),

$$\begin{aligned} f_2(v) &= \exp\left\{-\int_0^v \frac{\phi(b_n)}{\phi(a_n t + b_n)} dt\right\} \\ &= e^{-v}\Bigg\{1 + \frac{v^2}{2}\phi'(b_n) - \frac{v^3}{6}\{2\phi'(b_n)^2 - \phi(b_n)\phi''(b_n)\} + \frac{v^4}{8}\phi'(b_n)^2 \\ &\qquad + o((1+v^4)R(b_n))\Bigg\} \end{aligned} \tag{21}$$

where $R(b_n) = \{\phi'(b_n)\}^2 + |\phi(b_n)\phi''(b_n)|$, uniformly on $0 < v < -K \log\ |\phi'(b_n)|$.

Note that the argument in Smith (1987) primarily relies on a Taylor expansion of $\phi(a_n t + b_n)$. Although here the derivation is based on the Balkema and de Haan (1972) representation for $-\log F(x)$ as opposed to $1 - F(x)$, the Taylor expansion argument is the same; that is, the exact form of ϕ may be different but not the form of the Taylor expansion. Now all the steps of Smith's (1987) derivation of equation (9.20) apply also in the case $K \log|\varphi'(b_n)| < v < 0$ under the stronger assumptions (16) - (17). So (21) holds uniformly for $|v| < -K \log\ |\phi'(b_n)|$.

<u>*The expansion of f_3*</u> Since $f_3(v) = \frac{\phi(b_n)}{\phi(a_n v+b_n)}$ is embedded into the expansion of f_2, the expansion of f_3 falls from the condition (17) which implies

$$\phi'(a_n v + b_n)/\phi'(b_n) \to 1 \quad \text{and} \tag{22}$$

$$\phi(a_n v + b_n)/\phi(b_n) \to 1 \quad \text{as} \quad n \to \infty, \tag{23}$$

each uniformly over $|v| \leq e_n = -K \log\ |\phi'(b_n)|$.

In fact, using (16),(17), (22), and (23), and completing the Taylor's expansion

$$\begin{aligned} f_3(v) &= \frac{\phi(b_n)}{\phi(a_n v + b_n)} \\ &= 1 - v\phi'(b_n) + \frac{v^2}{2}\{2\phi'(b_n)^2 - \phi(b_n)\phi''(b_n)\} - \frac{v^3}{2}\phi'(b_n)^2 + o((1+v^4)R(b_n)) \end{aligned}$$

uniformly over $|v| \leq e_n = -K \log\ |\phi'(b_n)|$.

<u>*The expansion of f_1*</u> For $f_1(v) = \exp\ \{-\exp\ (\ -\int_0^v \frac{\phi(b_n)}{\phi(a_n t+b_n)}dt)\}$, we first use (17), (22), and (23) in the argument which led to the Smith (1987) equation above equation (9.19),

$$\int_0^v \frac{\phi(b_n)}{\phi(a_n t + b_n)}dt = v - \frac{v^2}{2}\phi'(b_n) + \frac{v^3}{6}\{2\phi'(b_n)^2 - \phi(b_n)\phi''(b_n)\} + o((1+|v|^3)R(b_n)))$$

uniformly over $|v| \leq e_n = -K \log\ |\phi'(b_n)|$.

Now write,

$$\int_0^v \frac{\phi(b_n)}{\phi(a_n t + b_n)}dt = v + T_n$$

where

$$T_n = T_n(v) = -\frac{v^2}{2}\phi'(b_n) + \frac{v^3}{6}\{2\phi'(b_n)^2 - \phi(b_n)\phi''(b_n)\} + o((1+|v|^3)R(b_n)).$$

Then $f_1(v) = \exp(\{-\exp\{-(v+T_n)\}\}$.

Expanding f_1 about v,

$$\begin{aligned} f_1(v) &= e^{-e^{-v}} + T_n e^{-v} e^{-e^{-v}} + \frac{T_n^2}{2} e^{-e^{-v}}(e^{-2v} - e^{-v}) \\ &\quad + \frac{T_n^3}{6}(1+o(1))e^{-e^{-v}}(e^{-3v} - 3e^{-2v} + e^{-v}) \\ &= e^{-e^{-v}}\Big\{1 - e^{-v}\frac{v^2}{2}\phi'(b_n) + \frac{v^3 e^{-v}}{6}[2\phi'(b_n)^2 - \phi(b_n)\phi''(b_n)] \\ &\quad + \frac{v^4(e^{-2v} - e^{-v})}{8}\phi'(b_n)^2 + o((1+e^{-2v})(1+v^4)R(b_n))\Big\} \end{aligned}$$

$$\forall |v| \leq -K \log|\phi'(b_n)|.$$

Together $f_{M_n^*}(v) = f_1 \times f_2 \times f_3$ Thus multiplying across and collecting terms gives $\forall |v| \le e_n$

$$f_{M_n^*}(v) = e^{-v}e^{-e^{-v}}\Big\{1+(\frac{v^2}{2}-v-\frac{v^2e^{-v}}{2})\phi'(b_n)+\{\frac{v^2}{2}+\frac{v^3}{6}[e^{-v}-1]\}\{(2\phi'(b_n)^2-\phi(b_n)\phi''(b_n)\}$$
$$+\{\frac{v^3}{2}[e^{-v}-2]+\frac{v^4}{8}[e^{-2v}-3e^{-v}+1]\}\phi'(b_n)^2+o((1+e^{-2v})(1+v^4)R(b_n))\Big\}. \tag{24}$$

Reconfiguring equation (24) into the form of equation (20) involves Lemma 9 with $m = 3$. Specifically, the constants that arise from equation (107) of Lemma 9 in the Appendix are absorbed into the ϵ_n^* function of equation (20). Thus the expansion in equation (20) holds $\forall |v| \le -K \log |\phi(b_n)|$.

To extend this expansion to the intervals $|v| \ge -K \log\ |\phi(b_n)|$, the idea is to show that all the terms in (20) are $o(|v|^{-\delta}\phi'(b_n)^2)$ for $|v| \ge e_n$ where $e_n = -K \log\ |\phi(b_n)|$.

Note it is sufficient to show the two terms $f_{M^*}(v)$ and $\Lambda'(v)$ are $o(|v|^{-\delta}\phi'(b_n)^2), |v| \ge e_n$ since the higher order terms associated with $\Lambda'(v)$ in (20) are of smaller order.

First look at Λ'. On the interval $|v| \ge e_n$, $\Lambda'(v)$ is maximized for sufficiently large n at $-e_n$ and e_n and hence the focus is on these two endpoints.

First, for $v = e_n$, since $K > 2$ it is possible to fix K' such that $1 < K' < K/2$. Thus (for $v > 0$)

$$|\Lambda'(v)| \le e^{-v} = e^{-\frac{v}{K'}}e^{-v(1-\frac{1}{K'})}.$$

But since exponential rates dominate polynomial,

$$e^{-v(1-\frac{1}{K'})} \le \kappa_1 |v|^{-\delta} \ \text{ for some constant } \kappa_1. \tag{25}$$

Now substitute in $v = e_n = -K \log |\phi'(b_n)|$,

$$e^{-\frac{v}{K'}} \le e^{\frac{K}{K'}\log|\phi'(b_n)|} = |\phi'(b_n)|^{\frac{K}{K'}} < |\phi'(b_n)|^2 \ \text{ since } \frac{K}{K'} > 2. \tag{26}$$

Combining equations (25) and (26), $|\Lambda'(v)| = o(|v|^{-\delta}\phi'(b_n))$ for $v = e_n$.

Second, for $v = -e_n$, again define K' as before

$$|\Lambda'(v)| < \frac{\kappa_2}{e^{-v}} = \kappa_2(e^{-v})^{-\frac{1}{K'}}(e^{-v})^{-1+\frac{1}{K'}} \ \text{ for some constant } \kappa_2.$$

Note in this case $v < 0$. Using the same argument in (25) and (26),

$$|\Lambda'(v)| = o(|v|^{-\delta}\phi'(b_n)) \text{ for } v = -e_n \text{ and hence for the entire interval } |v| \ge e_n. \tag{27}$$

Now (24) and (27) imply that $f_{M_n^*}(v) = o(|v|^{-\delta}\phi(b_n))$ for $v = \pm e_n$. Thus to prove (20) for $|v| \ge e_n$ it suffices to show that

(a.) $f_{M_n}(x)$ is increasing for $x \le -a_n e_n + b_n$,

(b.) $f_{M_n}(x)$ is decreasing for $x \ge a_n e_n + b_n$.

Taking the derivative of

$$f_{M_n}(x) = nf(x)F^{n-1}(x).$$

gives

$$\frac{d}{dx} f_{M_n}(x) = nF^{n-2}(x)\{f'(x)F(x) + (n-1)f^2(x)\}. \tag{28}$$

Solving the algebra,

$$(28) \begin{cases} \ge 0 & \text{if } \frac{f'(x)F(x)}{f^2(x)} \ge -(n-1) \\ \le 0 & \text{if } \frac{f'(x)F(x)}{f^2(x)} \le -(n-1). \end{cases} \tag{29}$$

Since (29) $\Rightarrow$ (20), we look at (29) .

Using (3) with $c(x) \equiv 1$

$$F(x) = \exp\left\{-\exp\{-\int_{-\infty}^{x} \frac{dt}{\phi(t)}\}\right\},$$

$$f(x) = \exp\left\{-\exp\{-\int_{-\infty}^{x} \frac{dt}{\phi(t)}\}\right\} \exp\left\{-\int_{-\infty}^{x} \frac{dt}{\phi(t)}\right\} \frac{1}{\phi(x)},$$

$$f'(x) = \frac{1}{\phi^2(x)} \exp\left\{-\exp\{-\int_{-\infty}^{x} \frac{dt}{\phi(t)}\}\right\} \exp\left\{-\int_{-\infty}^{x} \frac{dt}{\phi(t)}\right\} \times \left[\exp\left\{-\int_{-\infty}^{x} \frac{dt}{\phi(t)}\right\} + 1 + \phi'(x)\right].$$

Thus

$$\frac{-f'(x)F(x)}{f^2(x)} = 1 + \{1 + \phi'(x)\} \frac{1}{-\log F(x)}.$$

Using (3), $\phi'(x) \to 0$ as $x \to x_o$, so

$$\frac{-f'(x)F(x)}{f^2(x)} \sim \frac{1}{-\log F(x)}, \quad x \to x_o.$$

Hence, there exists some x^* such that, $\forall x \ge x^*$,

$$\frac{1}{2} \frac{1}{-\log F(x)} < \frac{-f'(x)F(x)}{f^2(x)} < \frac{2}{-\log F(x)}. \tag{30}$$

Now, by condition (16), $a_n e_n + b_n \ge x^*$ for all sufficiently large n, say $n \ge n_1$.

By condition (19), there exists a n_2 such that, whenever $n \ge n_2$

$$\inf_{x \le x^*} \frac{f'(x)F(x)}{f^2(x)} > -(n-1). \tag{31}$$

Let $n^* = \max(n_1, n_2)$. For $n \ge n^*, x^* \le x \le -a_n e_n + b_n$,

$$\frac{-f'(x)F(x)}{f^2(x)} < \frac{2}{-\log F(-a_n e_n + b_n)} \sim 2ne^{-e_n} < n - 1. \tag{32}$$

Note to simplify the denominator of (32)

$$-\log F(-a_n e_n + b_n) = -\log F(b_n)\left\{\frac{-\log F(-a_n e_n + b_n)}{-\log F(b_n)}\right\} \sim \frac{1}{n}e^{e_n}$$

seen by substituting $-e_n$ in for v in (103) and use the definition of b_n – namely, $-\log F(b_n) = \frac{1}{n}$.

To continue, putting (31) and (32) together, $\frac{f'(x)F(x)}{f^2(x)} > -(n-1)$ for the range $x \le -a_n e_n + b_n$.

For $x \ge a_n e_n + b_n$,

$$\frac{-f'(x)F(x)}{f^2(x)} > \frac{1}{2}\frac{1}{-\log F(a_n e_n + b_n)} \sim 2ne^{e_n} > n-1.$$

Thus for sufficiently large n,

$$\frac{-f'(x)F(x)}{f^2(x)}\begin{cases} \le n-1 & \text{if } x \le -a_n e_n + b_n \\ \ge n-1 & \text{if } x \ge a_n e_n + b_n. \end{cases}$$

This is equivalent to (29). Thus the result (20) holds. □

Corollary 2 *Given the notation and conditions in Proposition 1, then for each $\delta > 0$ there exists an n^* and a function ϵ_{n^*} tending to 0 as $n^* \to \infty$ such that $\forall n > n^*$ and $\forall |v| \le e_n = -K\log|\phi'(b_n)|$, and constant $j > 0$ finite,*

$$\begin{aligned}\Bigg| e^{-jv}f_{M_n^*}(v) - e^{-jv}\Lambda'(v)\Bigg\{1 + (\frac{v^2}{2} - v - \frac{v^2e^{-v}}{2})\phi'(b_n) \\ +(\frac{v^2}{2} + \frac{v^3}{6}[e^{-v}-1])(2\phi'(b_n)^2 - \phi(b_n)\phi''(b_n)) \\ (v^2 + \frac{5}{6}v^3[e^{-v}-1] + \frac{1}{8}v^4[e^{-2v} - 3e^{-v} + 1])\phi'(b_n)^2\Bigg\}\Bigg| \\ < \epsilon_{n^*}[\phi'(b_n)^2 + |\phi(b_n)\phi''(b_n)|\min(1,|v|^{-\delta})]. \qquad (33)\end{aligned}$$

Proof: This follows immediately from the proof of Proposition 1. For the interval $|v| \le e_n$, the proof is the same as in Proposition 1 except that $f_2(v)$ is replaced by $e^{-jv}f_2(v)$ to absorb the extra e^{-jv}. The changes are in (21) where

$$\begin{aligned} e^{-jv}f_2(v) &= e^{-(j+1)v}\Bigg\{1 + \frac{v^2}{2}\phi'(b_n) - \frac{v^3}{6}\{2\phi'(b_n)^2 - \phi(b_n)\phi''(b_n)\} + \frac{v^4}{8}\phi'(b_n)^2 \\ &\quad + o((1+e^{-2v})(1+v^4)R(b_n))\Bigg\}. \end{aligned}$$

The only effect this would have on the rate of convergence would be in (107) of Lemma 9 where m would now be $(j+3)$ but again this would be absorbed into the ϵ_{n^*} function.

Remark This gives us that $e^{-jv}f_{M_n^*}(v)$ is uniformly bounded on $|v| \le e_n$ for any $j > 0$ finite. □

3 Expansion for the conditional density of the sum given the maximum

Proposition 3 derives a higher order expansion for the conditional density of the sum given the maximum, specifically an expansion for the density of $f_{\tilde{S_n}}$, the sum normalized with respect to the conditional moments. Appendix B contains for lemmas and their corollaries used in deriving this proposition.

Proposition 3 *Assume f' is integrable, μ_3 exists, φ''' exists and is continuous in a neighborhood of 0, and $|\varphi_{u_n}(t)|^n$ is integrable for some $n \geq n^* > 1$. Then $f_{\tilde{S_n}}$ exists for $n \geq n^*$ and as $n \to \infty$*

$$f_{\tilde{S_n}}(x) - \mathcal{N}'(x) - \frac{\mathcal{K}_3(u_n)}{6\sigma^3(u_n)\sqrt{n}}(x^3 - 3x)\mathcal{N}'(x) = o(\frac{1}{\sqrt{n}}) \quad \text{uniformly in } x \text{ and } v \tag{34}$$

where the dependence on v is through the definition of u_n.

Proof: By Corollary 13 in Appendix B, the left hand side of (34) exists for $n \geq n^*$ and has Fourier norm

$$\begin{aligned} N_n &= \frac{1}{2\pi}\int_{-\infty}^{\infty}\left|\exp\Big(-\frac{it\mu(u_n)\sqrt{n-1}}{\sigma(u_n)}\Big)\varphi_{u_n}\Big(\frac{t}{\sqrt{n-1}\sigma(u_n)}\Big)^{n-1}\right. \\ &\qquad \left. - \exp(-\frac{t^2}{2}) - \frac{\mathcal{K}_3(u_n)}{6\sigma^3(u_n)\sqrt{n-1}}(it)^3\exp(-\frac{t^2}{2})\right|dt. \end{aligned}$$

Define N_L as N_n where the integral is restricted to the interval $|t| \leq \delta\sigma(u_n)\sqrt{n-1}$ and N_G as N_n where the integral is restricted to the intervals $|t| > \delta\sigma(u_n)\sqrt{n-1}$.

Now, choose $\delta > 0$ arbitrary but fixed. Substituting t by $\frac{t}{\sqrt{n-1}\sigma(u_n)}$ in equation (110) of Lemma 14, there exists a number $q_\delta < 1$ and a n^{**} such that

$$|\varphi_{u_n}(\frac{t}{\sqrt{n-1}\sigma(u_n)})| < q_\delta \quad \forall|t| > \delta\sigma(u_n)\sqrt{n-1} \text{ and } \forall n \geq n^{**}.$$

Thus $\forall n \geq n' = \max(n^*, n^{**})$, N_G is less than

$$q_\delta^{n-1-n'}\int_{-\infty}^{\infty}\left|\varphi_{u_n}(\frac{t}{\sqrt{n-1}\sigma(u_n)})\right|^{n'}dt + \int_{|t|>\delta\sigma(u_n)\sqrt{n-1}} e^{-\frac{t^2}{2}} \times \left\{1 + \left|\frac{\mathcal{K}_3(u_n)t^3}{6\sigma^3(u_n)\sqrt{n-1}}\right|\right\}dt. \tag{35}$$

Since $|\varphi_{u_n}(t)|$ is integrable for all $n \geq n' = max(n^*, n^{**})$ and q_δ does not depend on n, the first term of (35) tends to zero more rapidly than any power of $1/n$. The same holds for the second term and can be seen by substituting into the inequality $\inf_n \sigma(u_n)$ for $\sigma(u_n)$. Note since u_n is a threshold, $\sigma(u_n) > 0$. Thus $N_G = o(1/\sqrt{n})$ uniformly in n.

Substituting ψ_{u_n} from Lemma 15 into N_L,

$$N_L = \frac{1}{2\pi}\int_{|t|<\delta\sigma(u_n)\sqrt{n-1}} e^{-\frac{t^2}{2}}\left|\exp[(n-1)\psi_{u_n}(\frac{t}{\sqrt{n-1}\sigma(u_n)})] - 1 - \frac{\mathcal{K}_3(u_n)}{6\sigma^3(u_n)\sqrt{n-1}}(it)^3\right|dt$$

where $\psi_{u_n}(t) = \log \varphi_{u_n}(t) - it\mu(u_n) + \frac{t^2}{2}\sigma^2(u_n)$.

The integral will be evaluated using the following inequality from equation (2.9) of Feller (1971), p. 534,

$$|e^{\alpha} - 1 - \beta| \leq (|\alpha - \beta| + \frac{1}{2}\beta^2)e^{\gamma} \text{ where } \gamma = \max(|\alpha|, |\beta|).$$

Here

$$\alpha = (n-1)\psi_{u_n}(\frac{t}{\sqrt{n-1}\sigma(u_n)})$$

and

$$\beta = \frac{\mathcal{K}_3(u_n)}{6\sigma^3(u_n)\sqrt{n-1}}(it)^3 = (n-1)\frac{\mathcal{K}_3(u_n)}{6\sigma^3(u_n)(\sqrt{n-1})^3}(it)^3.$$

Thus

$$N_L \leq \frac{1}{2\pi}\int_{|t|<\delta\sigma(u_n)\sqrt{n-1}} e^{-t^2/2}|e^{\gamma}(\left|\alpha - \beta| + \frac{\beta^2}{2})\right| dt. \tag{36}$$

Substituting $\frac{t}{\sqrt{n-1}\sigma(u_n)}$ for t in equation (116) of Lemma 15,

$$\begin{aligned} |\alpha - \beta| &= (n-1)|\psi_{u_n}(\frac{t}{\sqrt{n-1}\sigma(u_n)}) - \frac{1}{6}\mathcal{K}_3(u_n)(\frac{it}{\sigma(u_n)(\sqrt{n-1})})^3| \\ &\leq (n-1)|\frac{t}{\sigma(u_n)\sqrt{n-1}}|^3\epsilon \\ &\leq \frac{\epsilon|t|^3}{\sqrt{n-1}\sigma(u_n)}. \end{aligned} \tag{37}$$

For the $\frac{1}{2}\beta^2$ term, algebraic simplification yields

$$\frac{1}{2}\beta^2 = \frac{t^6}{72(n-1)}\frac{\mathcal{K}_3^2(u_n)}{\sigma^6(u_n)}. \tag{38}$$

Note $\psi_{u_n}(t) \to \psi(t)$ as $n \to \infty$ uniformly in t by a series of dominated convergence theorem arguments. The moments $\kappa_3(u_n)$, $\sigma(u_n)$, and $\mu(u_n)$ are all uniformly bounded in n. Thus as $t \to 0$, both $|\alpha|$ and $|\beta|$ tend to 0 uniformly in n. In fact, it is possible to bound $|\alpha|$ and $|\beta|$. In particular, letting K_1 be a constant, we have

$$\begin{aligned} |\beta| &\leq \frac{K_1|t|^3}{\sqrt{n-1}\sigma^3(u_n)} \leq \frac{K_1|t|^2}{\sqrt{n-1}\sigma^3(u_n)}\delta\sigma(u_n)\sqrt{n-1} \\ &\leq \frac{K_1\delta|t|^2}{\sigma^2(u_n)} \\ &< \frac{|t|^2}{4} \text{ if we choose } \delta \text{ so that } \frac{K_1\delta}{\sigma(u_n)} < \frac{1}{4} \ \ \forall n > n^*. \end{aligned} \tag{39}$$

Similarly using the triangular inequality argument of Lemma 15 and defining K_2 as another constant,

$$
\begin{aligned}
|\alpha| &< K_2(n-1)|\frac{t}{\sqrt{n-1}\sigma(u_n)}|^3 \leq \frac{K_2 t^2}{\sqrt{n-1}\sigma^3(u_n)}\delta\sigma(u_n)\sqrt{n-1} \\
&\leq \frac{K_2\delta|t|^2}{\sigma^2(u_n)} \\
&< \frac{|t|^2}{4} \quad \text{if we choose } \delta \text{ so that } \tfrac{K_2\delta}{\sigma(u_n)} < \tfrac{1}{4} \quad \forall n > n^*. \qquad (40)
\end{aligned}
$$

Using (39) and (40),

$$\gamma < \frac{t^2}{4}. \qquad (41)$$

Using (37), (38), and (41), the integrand in (36) is less than

$$e^{\frac{-t^2}{4}}[\epsilon\frac{|t|^3}{\sqrt{n-1}\sigma(u_n)} + \frac{t^6}{72(n-1)}\frac{\mathcal{K}_3^2(u_n)}{\sigma^6(u_n)}].$$

Since ϵ is arbitrary and independent of x, $\sigma^2(u_n) \to \sigma^2$ and $\mathcal{K}_3(u_n) \to \mathcal{K}_3$ where σ^2 and $\mathcal{K}_3$ are assumed finite, and $\int_{-\infty}^{\infty} t^6 e^{-t^2} dt < \infty$, $N_L = o(1/\sqrt{n})$ uniformly in n. Thus (34) holds. □

4 Expansion of the Conditional Moments

Proposition 4 derives the expansions for the difference between the means and variances of X and X^* where the latter (conditional) random variable was defined in Section 1.4. The lemmas supporting this proposition are found in the Appendix C.

Proposition 4 *Under conditions of Proposition 1:*

$$\mu - \mu(u_n) \sim \begin{cases} \frac{u_n e^{-v}}{n} & x_o = \infty \\ \frac{(x_o-\mu)e^{-v}}{n} & x_o < \infty \end{cases} \qquad (42)$$

and

$$\sigma^2 - \sigma^2(u_n) \sim \begin{cases} \frac{u_n^2 e^{-v}}{n} & x_o = \infty \\ \{(x_o-\mu)^2 - \sigma^2\}\frac{e^{-v}}{n} & x_o < \infty \end{cases} \qquad (43)$$

uniformly on $|v| \leq e_n = -K\log|\phi'(b_n)|$.

If $x_o = \infty$ *and*

$$\frac{\phi'(t)\log|\phi'(t)|}{t} \to 0, \quad \textit{as } t \to \infty, \qquad (44)$$

then

$$\mu - \mu(u_n) \sim \frac{b_n e^{-v}}{n} \qquad (45)$$

$$\sigma^2 - \sigma^2(u_n) \sim \frac{b_n^2 e^{-v}}{n} \qquad (46)$$

uniformly in $|v| \leq e_n = -K\log|\phi'(b_n)|$.

Proof Recall

$$E(X|X < u_n) = \mu(u_n), E(X|X \geq u_n) = u_n + m(u_n),$$

$$Var(X|X < u_n) = \sigma^2(u_n), Var(X|X \geq u_n) = s^2(u_n).$$

For the proof of (42), rearranging a law of total probability yields

$$\mu - \mu(u_n) = \{\frac{1 - F(u_n)}{F(u_n)}\}\{u_n + m(u_n) - \mu\}. \tag{47}$$

To determine (42), we expand and compare the terms in the RHS of (47), in particular, the terms $F(u_n)$, $m(u_n)$, and $1 - F(u_n)$. First, $F(u_n) \to 1$ as $n \to \infty$ and hence has no influence on the results.

Second, under conditions similar to Proposition 1, Section 9.2 of Smith(1987) establishes the expansion for the mean of the exceedance:

$$E[X - u_n|X > u_n] = m(u_n) = \phi(u_n) + \phi(u_n)\phi'(u_n) + o(|\phi(u_n)\phi'(u_n)|) \quad \forall |v| \leq e_n. \tag{48}$$

where $e_n = -K \log |\phi'(b_n)|$.

To see that this term does not contribute to (42), we look at two cases:

When $x_o < \infty$: Using (127) of Lemma 16 shows that $\phi(u_n) \to 0$ as $n \to \infty$ and thus does not contribute.

When $x_o = \infty$: From(126) of Lemma 16, we have that $\frac{\phi(u_n)}{u_n} \to 0$ as $n \to \infty$. Thus in (47) the term u_n contributes while $m(u_n) \sim \phi(u_n)$ does not.

Finally, rewriting $F(u_n)$ and taking a Taylor expansion of e^x yields

$$\begin{aligned} 1 - F(u_n) &= 1 - e^{\log F(u_n)} \\ &= 1 - \{1 + \log F(u_n) + (1/2 + o(1)) \log^2 F(u_n)\} \\ &= -\log F(u_n)\{1 + (1/2 + o(1)) \log F(u_n)\}. \end{aligned} \tag{49}$$

From (103) of Lemma 8 and using the definition of b_n in (4)

$$-\log F(u_n) = \frac{e^{-v}}{n}[1 + O(v^2|\phi'(b_n)|)], \quad \text{uniformly on } |v| \leq e_n. \tag{50}$$

Putting (49) and (50) together

$$1 - F(u_n) = \frac{e^{-v}}{n}(1 + o(1)), \quad \text{uniformly on } |v| \leq e_n. \tag{51}$$

This term does contribute to the higher order term in (42) which is obtained by multiplying (51) to the term u_n in (47).

A similar derivation exists for (43): the difference between the unconditional and conditional variance. Rearranging the law of total probability in this case gives

$$\begin{aligned} \sigma^2 - \sigma^2(u) &= \sigma^2 + \mu^2(u) - \{\frac{\mu^2+\sigma^2}{F(u)}\} + \{\frac{1}{F(u)} - 1\}\{(u+m(u))^2 + s^2(u)\} \\ &= \mu^2(u) - \mu^2 + \{\frac{1-F(u)}{F(u)}\}\{(u+m(u))^2 + s^2(u) - \mu^2 - \sigma^2\}. \end{aligned} \quad (52)$$

Again concluding (43) involves comparing the expansions for the terms of the RHS of (52). For the first two terms, the expansion needed is obtained by squaring (42) and retaining the leading terms:

$$\mu^2 - \mu^2(u_n) \sim \begin{cases} 2\mu u_n \frac{e^{-v}}{n} & x_o = \infty \\ 2(x_o - \mu)\mu\frac{e^{-v}}{n} & x_o < \infty, \quad \forall |v| \le e_n. \end{cases} \quad (53)$$

The only remaining term in (52) to expand is $s^2(u_n)$. Under conditions similar to Proposition 1, Section 9.2 of Smith (1987) establishes

$$E\{Y^2|X > u_n\} = 2\phi^2(u_n) + 6\phi^2(u_n)\phi'(u_n) + o(\phi^2(u_n)(\phi'(u_n))^2), \quad forall|v| \le e_n.$$

from which we obtained with (48), $forall|v| \le e_n$,

$$s^2(u_n) = E\{Y^2|X > u_n\} - E^2\{Y|X > u_n\} = \phi^2(u_n) + 4\phi^2(u_n)\phi'(u_n) + o(\phi^2(u_n)\phi'(u_n)). \quad (54)$$

Using the relationship in (126) and (127) of Lemma 16 and substituting the expansions (48), (54), (53), and (51) into (52), the higher order terms fall from the $1 - F(u_n)$ and the $\{u_n + m(u_n)\}^2$ terms. Specifically the comparison of the terms yields (43).

Remark (1.) We note the following differences between the conditions in Proposition 1 and Proposition 9.2 of Smith (1987). First, Smith (1987) assumes the Balkema and de Haan (1972) representation for $1 - F(x)$ as opposed to $-\log F(x)$ which used in Proposition 1. In this application, the difference between $-\log F$ and $1 - F$ is $o(\frac{1}{n})$ which is of smaller order than the leading terms in these expansions. Thus this difference does not impact the results. Second, Smith(1987) assumes $c(u) - 1 \sim s\{(\phi'(u))^2 + |\phi(u)\phi''(u)|\} \to 0$ as $u \to x_o$ for finite s. Here the assumption is $c(x) \equiv 1$. Again the remainder term for c in Smith(1987) is of smaller order than the leading terms obtained here and therefore assuming $c(x) \equiv 1$ does not impact the validity of the above formula for this application.

Remark (2.) Now to establish the case when $x_o = \infty$, the added assumption (44) is necessary to guarantee that we can replace u_n by b_n [i.e. $\frac{u_n}{b_n} \to 1$] uniformly on $|v| \le e_n$. This condition results from the following assumptions in Lemma 1 of Cohen(1982a): [A] Either (1) $\phi'(u) > 0$ for all sufficiently large u or (2) $\phi'(u) < 0$ for all sufficiently large u ***and*** [B] $\phi'(u)$ is regularly varying for $u \to \infty$.

Given [A] and [B], then (44) falls from formula (a) at the bottom of page 846 of Cohen(1982a). Specifically assumption (44) allows the following on the interval $|v| \leq e_n = -K \log|\phi(b_n)|$ (recall $\phi(b_n) = a_n$)

$$\begin{aligned} \frac{u_n}{b_n} = \frac{a_n v + b_n}{b_n} &\leq \frac{\phi(b_n)\{-K \log|\phi'(b_n)|\} + b_n}{b_n} \\ &= 1 + -K \frac{\phi(b_n) \log|\phi'(b_n)|}{b_n} \\ &\rightarrow 1 \text{ as } n \rightarrow \infty, \text{ uniformly in } |v| \leq e_n. \end{aligned}$$

Thus $u_n = b_n(1 + o(1))$, $|v| \leq e_n$ and allows u_n to be replaced by b_n in (42) and (43) without changing the results leading to (45) and (46).

Remark (3.) Finally, in establishing (45) and (46) it is necessary that $\frac{b_n^2}{n} \rightarrow 0$. To see this, rewrite

$$\frac{b_n^2}{n} = b_n^2\{-\log F(b_n)\} = b_n^2\{1 - F(b_n) + o(\frac{1}{n})\}.$$

Now

$$b_n^2\{1 - F(b_n)\} \leq \int_{b_n}^{x_o} x^2 dF(x) = \int_{-\infty}^{\infty} 1_{(b_n, x_o)} x^2 dF(x).$$

Since variance is assumed to be finite, then by dominated convergence theorem

$$b_n^2\{1 - F(b_n)\} \rightarrow 0 \text{ as } n \rightarrow \infty.$$

Thus

$$\frac{b_n^2}{n} \rightarrow 0 \text{ as } n \rightarrow \infty, \quad \forall |v| \leq e_n. \tag{55}$$

□

5 Expansion of the Joint Density

This section presents the main result and its corollaries with their respective proofs. Specifically, Theorem 5 provides the main result: an expansion for the joint density of the sum and the maximum. Corollary 6 is model oriented in that it substitutes the limiting densities for $f_{M_n^*}$ and $f_{S_n^*}$ into Theorem 5. Finally, corollary 7 is the penultimate version of Theorem 5.

Theorem 5 *Let $X_1, \ldots, X_n$ be an iid sequence of random variables with distribution function F, density function f, characteristic function φ, mean μ, and variance σ^2. Let u_n be a threshold level and φ_{u_n} be the characteristic function of the random variable $X|X < u_n$.*

We make the following two sets of assumptions:

Set A: Assume f' is integrable, μ_3 exists, φ''' exists and is continuous in a neighborhood of 0, and $|\varphi_{u_n}(t)|^n$ is integrable for $n \geq$ some $n^* > 1$.

Set B: Assume F is in the domain of attraction of Λ so that the representation in (3) holds. Use the same form of a_n and b_n as defined in (5) and (4). Also assume that for a constant $K > 2$,

$$a_n v + b_n \to x_o \quad \textit{uniformly in } |v| \le -K \log\ |\phi'(b_n)| \textit{ as } n \to \infty \tag{56}$$

$$\phi''(a_n v + b_n)/\phi''(b_n) \to 1 \quad \textit{uniformly in } |v| \le e_n = -K \log\ |\phi'(b_n)| \textit{ as } n \to \infty \tag{57}$$

$$\phi(b_n)\phi''(b_n) \log\ |\phi'(b_n)|/\phi'(b_n) \to 0 \quad \textit{as } n \to \infty. \tag{58}$$

In addition, we assume that for any $v^* < x_o$

$$\inf_{v \le v^*} \frac{f'(v)F(v)}{f^2(v)} > -\infty \tag{59}$$

and when $x_o = \infty$,

$$\frac{\phi'(t) \log |\phi'(t)|}{t} \to 0, \quad \textit{as } t \to \infty, \tag{60}$$

Define

$$r_n = \begin{cases} \frac{b_n}{\sqrt{n\sigma^2}}, & x_o = \infty, \\ \frac{x_o - \mu}{\sqrt{n\sigma^2}} & x_o < \infty. \end{cases} \tag{61}$$

Then

$$|f_{S_n^*, M_n^*}(w, v) - f_{S_n^*}(w) f_{M_n^*}(v)\{1 - r_n(e^{-v} - 1)w\}| = o(r_n) \tag{62}$$

uniformly $\forall w$ and $\forall |v| \le -K \log |\phi'(b_n)|$.

Corollary 6 *Given the conditions in Theorem 5, if $x_o = \infty$*

$$\left| f_{S_n^*, M_n^*}(w, v) - \mathcal{N}'(w)\Lambda'(v)\{1 + (\frac{v^2}{2} - v - \frac{v^2 e^{-v}}{2})\phi'(b_n)\}\{1 - r_n(e^{-v} - 1)w\} \right| = o\{\max(r_n, |\phi'(b_n)|)\} \tag{63}$$

and if $x_o < \infty$

$$\left| f_{S_n^*, M_n^*}(w, v) - \mathcal{N}'(w)\Lambda'(v)\{1 + \frac{\kappa_3}{6\sigma^3\sqrt{n}}(w^3 - 3w)\} \times \{1 + (\frac{v^2}{2} - v - \frac{v^2 e^{-v}}{2})\phi'(b_n)\}\{1 - r_n(e^{-v} - 1)w\} \right| = o\{\max(r_n, |\phi'(b_n)|)\} \tag{64}$$

uniformly $\forall w$ and $|v| \le -K \log |\phi'(b_n)|$.

Corollary 7 *Given the conditions in Theorem 5 and defining*

$$H(x;\eta,\psi,k) = \exp\left[-\{1 - \frac{k(x-\eta)}{\psi}\}_+^{1/k}\right],$$

let $k_n = -\phi'(b_n)$ *and replace*

$$\Lambda'(v)\{1 + (\frac{v^2}{2} - v - \frac{v^2 e^{-v}}{2})\phi'(b_n)\} \tag{65}$$

in (63) and (64) by

$$H'(v;0,1,k_n) \quad \text{where} \quad H'(x;\eta,\psi,k) = \frac{d}{dx}H(x;\eta,\psi,k),$$

then we obtain the same result as Theorem 5.

Proof of Theorem 5

The joint density of S_n^* and M_n^* can be written as

$$\begin{aligned} f_{S_n^*,M_n^*}(w,v) &= f_{M_n^*}(v) f_{S_n^*|M_n^*}(w|v) \\ &= f_{M_n^*}(v)\sqrt{\frac{n\sigma^2}{(n-1)\sigma^2(u_n)}} f_{\tilde{S_n}}(z) \end{aligned}$$

where

$$z = \frac{n\mu + \sqrt{n\sigma^2}w - (n-1)\mu(u_n) - u_n}{\sqrt{(n-1)\sigma^2(u_n)}}. \tag{66}$$

.

Here $u_n = a_n v + b_n$. To enable the uniformity results v, w and hence z are allowed to be dependent on n. In general, this dependence on n is suppressed so as to make the notation easier to read. Note the transformation from S_n^* to $\tilde{S_n}$ and the form of (66) comes from (6).

To establishing (62) the first step is to break its LHS up as follows

$$f_{M_n^*}(v)\sqrt{\frac{n\sigma^2}{(n-1)\sigma^2(u_n)}} f_{\tilde{S_n}}(z) - f_{S_n^*}(w) f_{M_n^*}(v)\{1 - r_n(e^{-v}-1)w\} = E_1+E_2+E_3+E_4+E_5+E_6 \tag{67}$$

where

$$\begin{aligned}
E_1 &= f_{M_n^*}(v) f_{\tilde{S_n}}(z)\left[\sqrt{\frac{n\sigma^2}{(n-1)\sigma^2(u_n)}} - 1\right] && (68)\\
E_2 &= f_{M_n^*}(v)\left[f_{\tilde{S_n}}(z) - \mathcal{N}'(z)\{1 + \frac{\kappa_3(u_n)}{6\sigma^3(u_n)\sqrt{n}}(z^3 - 3z)\}\right] && (69)\\
E_3 &= f_{M_n^*}(v)\left[\mathcal{N}'(z) - \mathcal{N}'(w)\{1 - r_n(e^{-v}-1)w\}\right] && (70)\\
E_4 &= f_{M_n^*}(v)\left[\mathcal{N}'(w)\{1 + \frac{\kappa_3}{6\sigma^3\sqrt{n}}(w^3 - 3w)\} - f_{S_n^*}(w)\right]\{1 - r_n(e^{-v}-1)w\} && (71)\\
E_5 &= f_{M_n^*}(v)\left[\frac{\mathcal{N}'(z)\kappa_3(u_n)}{6\sigma^3(u_n)\sqrt{n}}(z^3 - 3z) - \frac{\mathcal{N}'(w)\kappa_3}{6\sigma^3\sqrt{n}}(w^3 - 3w)\right] && (72)\\
E_6 &= f_{M_n^*}(v)\mathcal{N}'(w)\frac{\kappa_3}{6\sigma^3\sqrt{n}}(w^3 - 3w)[r_n(e^{-v}-1)w]. && (73)
\end{aligned}$$

To prove (62), it suffices to prove that for $j = 1, 2, ..., 6$ that $E_j = o(r_n)$ uniformly $\forall\ v \leq e_n$ and $\forall w$.

Proof for E_1. The form of the higher order term in E_1 – (68) – falls from its third term.

Using (43),

$$\frac{\sigma^2(u_n)}{\sigma^2} = \begin{cases} 1 - \frac{b_n^2 e^{-v}}{n} + o(\frac{b_n^2 e^{-v}}{n}) & x_o = \infty \\ 1 - \{(x_o - \mu)^2 - \sigma^2\}\frac{e^{-v}}{n} + o[\{(x_o - \mu)^2 - \sigma^2\}\frac{e^{-v}}{n}] & x_o < \infty. \end{cases}$$

Inverting the above formula yields

$$\frac{\sigma^2}{\sigma^2(u_n)} = \begin{cases} 1 + \frac{b_n^2 e^{-v}}{n} + o(\frac{b_n^2 e^{-v}}{n}) & x_o = \infty \\ 1 + \{(x_o - \mu)^2 - \sigma^2\}\frac{e^{-v}}{n} + o[\{(x_o - \mu)^2 - \sigma^2\}\frac{e^{-v}}{n}] & x_o < \infty. \end{cases} \tag{74}$$

This inversion holds as long as $\frac{b_n^2 e^{-v}}{n} \to 0$ uniformly on $|v| \leq e_n$ which is true by (55).

Given $\sqrt{\frac{n}{n-1}} = 1 + O(\frac{1}{n})$ and $\frac{\sigma^2}{\sigma^2(u_n)} \geq 1$ and bounded,

$$\left| \sqrt{\frac{n\sigma^2}{(n-1)\sigma^2(u_n)}} - 1 \right| \leq \left\{ \frac{\sigma^2}{\sigma^2(u_n)} - 1 \right\} + O(\frac{1}{n}). \tag{75}$$

Substituting (74) into (75), for some κ

$$\left| \sqrt{\frac{n\sigma^2}{(n-1)\sigma^2(u_n)}} - 1 \right| \leq \begin{cases} \kappa(\frac{b_n^2 e^{-v}}{n} + \frac{1}{n}) & x_o = \infty \\ \kappa(\frac{e^{-v}}{n} + \frac{1}{n}) & x_o < \infty. \end{cases} \tag{76}$$

Using the r_n notation defined in (61) in (76),

$$E_1 = O(e^{-v} f_{M_n^*}(v) f_{\tilde{S_n}}(z) r_n^2).$$

Proposition 3 gives $f_{\tilde{S_n}}(z)$ is bounded $\forall z$. Corollary 2 gives $e^{-v} f_{M_n^*}(v)$ on $|v| \leq e_n$. Therefore

$$E_1 = o(r_n) \text{ for } |v| \leq e_n \text{ and } \forall w. \tag{77}$$

Proof for E_2. In (69), $\left[f_{\tilde{S_n}}(z) - \mathcal{N}'(z)\{1 + \frac{\kappa_3(u_n)}{6\sigma^3(u_n)\sqrt{n}}(z^3 - 3z)\} \right] = o(\frac{1}{\sqrt{n}})$ uniformly in z by Proposition 3 and $f_{M_n^*}(v)$ is bounded $\forall v$ by Proposition 1. Hence

$$E_2 = o(\frac{1}{\sqrt{n}}) = o(r_n), \quad \forall w, \forall v. \tag{78}$$

Proof for E_3. Although more involved, the key to resolving E_3 in (70) is first to take a Taylor's expansion of $\mathcal{N}'(z)$ about $\mathcal{N}'(w)$. Once we substituting the result of the Taylor's expansion back into E_3, the argument then focuses on the difference between z and w. At that point – Step 4 – we consider two cases where the dependence on n for z, w, and v need to be explicitly expressed: Case (a) where $|z_n - w_n|$ is bounded above and Case (b) where $|z_n - w_n|$ is bounded away from 0.

Step 1: Rearranging (66) provides an explicit form of the difference between z and w:

$$z - w = \frac{n(\mu - \mu(u_n))}{\sqrt{(n-1)\sigma^2(u_n)}} + \left\{ \sqrt{\frac{n\sigma^2}{(n-1)\sigma^2(u_n)}} - 1 \right\} w + \frac{\mu(u_n) - u_n}{\sqrt{(n-1)\sigma^2(u_n)}}.$$

For $x_o = \infty$, use (45), (46), and (76) and simplify to

$$z - w = \frac{b_n e^{-v}}{\sqrt{n\sigma^2}}\{1 + o(1)\} + O(we^{-v}(\frac{b_n}{\sqrt{n\sigma^2}})^2) - \frac{b_n}{\sqrt{n\sigma^2}}\{1 + o(1)\}. \tag{79}$$

For $x_o < \infty$, use (42), (43), and (76) and simplify to

$$z - w = \frac{(x_o - \mu)e^{-v}}{\sqrt{n\sigma^2}}\{1 + o(1)\} + O(w(1 + e^{-v})(\frac{b_n}{\sqrt{n\sigma^2}})^2) + \frac{(\mu - x_o)(1 + o(1))}{\sqrt{n\sigma^2}}. \tag{80}$$

Using the definition of r_n from (61), combine (79) and (80)

$$z - w = r_n(e^{-v} - 1) + o\{r_n(e^{-v} + 1)\} + O(w(e^{-v} + 1)r_n^2). \tag{81}$$

Step 2: Take a Taylor expansion for $\mathcal{N}'(z)$ about $\mathcal{N}'(w)$.

Using $z = w + t_n$ where t_n is seen in (81), write

$$\mathcal{N}'(z) = \mathcal{N}'(w + t_n) = \mathcal{N}'(w) + t_n\mathcal{N}''(z^*) \quad \text{where} \quad z^* \text{ between } w \text{ and } z.$$

Using the identity $\mathcal{N}''(x) = -x\mathcal{N}'(x)$,

$$\mathcal{N}'(z) = \mathcal{N}'(w) - t_n z^*\mathcal{N}'(z^*) = \mathcal{N}'(w) - (z - w)\mathcal{N}'(z^*)$$

for z^* between w and z.

Substituting this into E_3 and adding and subtracting a $f_{M_n^*}(v)(z - w)w$ term

$$\begin{aligned} E_3 &= f_{M_n^*}(v)(z - w)\left[w\mathcal{N}'(w) - z^*\mathcal{N}'(z^*)\right] + f_{M_n^*}(v)\mathcal{N}'(w)r_n(e^{-v} - 1)w \\ &\quad - f_{M_n^*}(v)(z - w)w\mathcal{N}'(w) \\ &= f_{M_n^*}(v)(z - w)\left[w\mathcal{N}'(w) - z^*\mathcal{N}'(z^*)\right] + o(r_n(e^{-v} + 1)w\mathcal{N}'(w)f_{M_n^*}(v)) \\ &\quad + O(w(e^{-v} + 1)r_n^2 w\mathcal{N}'(w)f_{M_n^*}(v)) \\ &= E_7 + E_8 + E_9. \end{aligned}$$

Step 3: Standard properties of the normal density shows $w^2\mathcal{N}'(w)$ is bounded. By Corollary 2, $(1 + e^{-v})f_{M_n^*}(v)$ is bounded. Thus

$$E_8 = o(r_n) \quad \text{uniformly } \forall w \text{ and } |v| \le e_n \tag{82}$$

and

$$E_9 = O(r_n^2) \quad \text{uniformly } \forall w \text{ and } |v| \le e_n. \tag{83}$$

Step 4: Now we focus on E_7. The important details in this formula concern $z - w$ since the other terms are bounded. Recall $v, w, z,$ and z^* actually depend on n. So fix the notation by writing $v = v_n$, $z = z_n$, $w = w_n$, and $z^* = z_n^*$ so the dependence on n is explicit.

Substituting (81) into E_7 and adding and subtracting a $z_n^{*2}\mathcal{N}'(z_n^*)$ term yields

$$\begin{aligned}E_7 &= f_{M_n^*}(v_n)r_n(e^{-v_n}-1)[w_n\mathcal{N}'(w_n)-z_n^*\mathcal{N}'(z_n^*)] \\ &\quad +o\{f_{M_n^*}(v_n)r_n(e^{-v_n}+1)[w_n\mathcal{N}'(w_n)-z_n^*\mathcal{N}'(z_n^*)]\} \\ &\quad +O(f_{M_n^*}(v_n)r_ne^{-v_n}[w_n^{*2}\mathcal{N}'(w_n)-z_n^{*2}\mathcal{N}'(z_n^*)]) \\ &\quad +O(f_{M_n^*}(v_n)e^{-v_n}r_n^2(w_n-z_n^*)z_n^*\mathcal{N}'(z_n^*)) \\ &= E_{10}+E_{11}+E_{12}+E_{13}.\end{aligned}$$

At this point, it is necessary to separate the argument into the two cases.

Case (a). By uniform continuity of $w_n^k\mathcal{N}'(w_n)$ for $k=0,1,2$ given $\epsilon>0$ there exists a $\delta>0$ such that

$$|z_n-w_n|<\delta \Rightarrow |z_n^*-w_n|<\delta \Rightarrow |z_n^{*k}\mathcal{N}'(z_n)-w_n^k\mathcal{N}'(w_n)|<\frac{\epsilon}{C}$$

for $k=0,1,2$ and any given constant $C>0$.

Now since $(e^{-v_n}+1)f_{M_n^*}(v_n)$ is bounded on $|v_n|\le e_n$, each E_{10}, E_{11}, and E_{12} is bounded by some

$$\textit{constant}\times|z_n^{*k}\mathcal{N}'(z_n^*)-w_n^k\mathcal{N}'(w_n)|$$

for $k=0,1,2$.

In other words, it is possible to choose a δ so that $\forall w_n$ and $\forall|v_n|\le e_n$

$$|z_n-w_n|<\delta \Rightarrow E_{10}=E_{11}=E_{12}=o(r_n). \tag{84}$$

As for E_{13}, $f_{M_n^*}(v_n)e^{-v}$ is bounded on $|v|\le e_n$, $z_n^*\mathcal{N}'(z_n^*)$ is bounded by standard normal density properties, and here $(z_n^*-w_n)\le\delta$, thus

$$E_{13}=O(r_n^2)=o(r_n), \forall v\le e_n. \tag{85}$$

Together (82), (83),(84) and (85) show that

$$E_3=o(r_n) \ \text{ uniformly } |v_n|\le e_n \text{ and } \forall w_n \tag{86}$$

when $|z_n-w_n|\le\delta$. This establishes Case (a).

Case (b). The premise behind Case (b) is that if $|z_n-w_n|>\tau$ – for any $\tau>0$ – then the entire left-hand side of (67) is $o(r_n)$. Once that is established the necessary result falls immediately since if the above result hold for any τ than it holds for the δ in Case (a).

Part 1 Suppose $|z_n-w_n|>\tau$ and $r_ne^{-v_n}<\tau^2$.

If $r_n e^{-v_n} < \tau$, then a consequence of (81) is that $|w_n r_n| >$ some $\tau_1 > 0$ for all sufficiently large n. In other words, $|w_n| > \frac{\tau_1}{r_n}$. Also from (81),

$$\begin{aligned} z_n &= w_n + r_n(e^{-v_n} - 1) + o(r_n(e^{-v_n} + 1)) + O(w_n(e^{-v_n} + 1)r_n^2) \\ &= w_n + o(1) + o(1) + O(w_n o(1)) \\ &= w_n(1 + o(1)) \\ &> \frac{\tau_1}{2r_n}, \quad \text{say,} \end{aligned}$$

for all sufficiently large n.

This relationship $z_n > \frac{\tau_1}{2r_n}$ [with (76)] implies

1. $|f_{\tilde{S_n}}(z_n)| = o(r_n)$ by Proposition 3 and
2. $|w_n|^k |f_{S_n^*}(w_n)| = o(r_n)$ for k=0,1 by Theorem 17 of Petrov(1975).
3. Note again by Proposition 1 and Corollary 2, $|(e^{-v_n} + 1)f_{M_n^*}(v_n)|$ is bounded on $|v_n| \leq e_n$.

Hence the left-hand side of (67) is $o(r_n)$, $\forall |v_n| \leq e_n$ and $\forall w_n$ when $|z_n - w_n| > \tau$ and $r_n e^{-v_n} < \tau^2$.

Part 2 Suppose $|z_n - w_n| > \tau$ and $r_n e^{-v} \geq \tau^2$. Corollary 2 says that for any finite $m' > 2$, $e^{-m' v_n} e^{-v_n} f_{M_n^*}(v_n)$ is uniformly bounded on $|v_n| \leq e_n$. This gives $e^{-v_n} f_{M_n^*}(v_n) = O(e^{(m'-1)v_n})$ but if $r_n e^{-v_n} \geq \tau^2 > 0$ then $r_n e^{-v_n} \not\to 0$ so $|v_n| < \log r_n$ for sufficiently large n. Thus

$$e^{-v_n} f_{M_n^*}(v_n) = O(e^{(m'-1)v_n}) \leq O(r_n^{(m'-1)}) = o(r_n) \text{ since } m' \text{ can be taken} > 2. \tag{87}$$

Now looking at the parts of (67), first $f_{\tilde{S_n}}(z_n)$ is bounded – see Proposition 3 – and second $w_n f_{S_n^*}(w_n)$ is bounded – again, see Theorem 17 of Petrov (1975). Finally from (87) $e^{-v} f_{M_n^*}(v_n) = o(r_n)$ which also gives $f_{M_n^*}(v_n) = o(r_n)$. This establishes Case (b).

Taking Part 1 and Part 2 together, the left-hand side of (67) is $o(r_n)$ when $|z_n - w_n| > \tau$, including $\tau = \delta$.

Thus for E_3 either

$$E_3 = o(r_n) \tag{88}$$

or the entire left-hand side of (67) is $o(r_n)$, uniformly on $|v| \leq e_n$ and $\forall w$.

Proof of E_4: In (71), again $f_{M_n^*}(v)$ is bounded and also by the unconditional Edgeworth expansion theorem – Feller (1971), Chapter XVI, Section 2, Theorem 1 – the term inside $[\mathcal{N}'(w)\{1 + \frac{\kappa_3}{6\sigma^3\sqrt{n}}(w^3 - 3w)\} - f_{S_n^*}(w)]$ is $o(\frac{1}{\sqrt{n}})$ uniformly in w.

Steps (a) – (c) show how to deal with multiplying by the term $r_n(e_{-v} - 1)w_n$:

(a) Corollary 2 gives that $\sup_v (e^{-v} - 1) f_{M_n^*}(v)$ is bounded on $|v| \leq e_n$.

(b) Theorem 17 of Petrov (1975) central limit theorem [which hold under the same assumptions as Proposition 3] gives $\sup_w |w\{\mathcal{N}'(w) - f_{S_n^*}(w)\}| \to 0$ as $n \to \infty$.

(c) Standard properties of the normal density imply $w^4\mathcal{N}'(w)$ is bounded so $\mathcal{N}'(w)\frac{\kappa_3}{6\sigma^3\sqrt{n}}(w^3 - 3w)w \to 0$ uniformly in w.

Thus

$$|E_4| = o(r_n) \;\text{ for } |v| \le e_n \text{ and } \forall w. \tag{89}$$

Proof of E_5: In (72), again, $f_{M_n^*}(v)$ is uniformly bounded on $|v_n| \le e_n$. Also the function $\mathcal{N}'(z)\{z^3 - 3z\}$ is uniformly continuous so by similar argument to proof of E_3, particularly *Step 3*,

$$E_5 = o(r_n) \;\; \forall |v_n| \le e_n, \forall w. \tag{90}$$

Proof of E_6: Now, (73) falls immediately since,

(a) $\sup_v (e^{-v} - 1) f_{M_n^*}(v)$ is bounded on $|v| \le e_n$ by Corollary 2.

(b) $w^4\mathcal{N}'(w)$ is bounded uniformly in w so $\mathcal{N}'(w)\frac{\kappa_3}{6\sigma^3}(w^3 - 3w)w$ is bounded uniformly in w by properties of the normal density.

Thus

$$E_6 = O(\frac{r_n}{\sqrt{n}}) = o(r_n) \;\text{ for } |v| \le e_n \text{ and } \forall w. \tag{91}$$

In conclusion, (77),(78), (88), (89), (90), and (91) shows that $E_j = o(r_n)$, for $j = 1, 2, 3, 4, 5, 6$ uniformly $\forall w$ and $|v| \le e_n = -K \log |\phi'(b_n)|$. □

Proof of Corollary 6

First we prove (64) under the assumption $x_o < \infty$.

Substituting the result of Theorem 5, equation (62), into (64) the form of the result to be proven is

$$\begin{aligned} \Bigg| \Bigg[f_{S_n^*}(w) f_{M_n^*}(v) - \mathcal{N}'(w)\{1 + \frac{\kappa_3}{6\sigma^3\sqrt{n}}(w^3 - 3w)\}\Lambda'(v) \\ \times\{1 + (\frac{v^2}{2} - v - \frac{v^2 e^{-v}}{2})\phi'(b_n)\}\Bigg] \{1 - r_n(e^{-v} - 1)w\} \Bigg| \\ = o(\max\{r_n, |\phi'(b_n)|\}) \end{aligned} \tag{92}$$

for $\forall w$ and $\forall |v| \le e_n = -K \log |\phi'(b_n)|$.

First we show (92) without the $r_n(e^{-v} - 1)w$ term.

Let

$$A_n = f_{S_n^*}(w),$$

$$A'_n = \mathcal{N}'(w)\{1 + \frac{\kappa_3}{6\sigma^3\sqrt{n}}(w^3 - 3w)\},$$

$$B_n = f_{M_n^*}(v),$$

and

$$B_n' = \Lambda'(v)\{1 + (\frac{v^2}{2} - v - \frac{v^2 e^{-v}}{2})\phi'(b_n)\}.$$

Then the left-hand side of (92) without the $r_n(e^{-v} - 1)w$ term can be written as

$$\begin{aligned} |A_n B_n - A_n' B_n'| &= |A_n B_n - A_n B_n' + A_n B_n' - A_n' B_n'| \\ &\leq |A_n|\,|B_n - B_n'| + |B_n'|\,|A_n - A_n'|. \end{aligned} \tag{93}$$

Now $|A_n|$ is bounded $\forall w$ by Theorem 17 of Petrov (1975) and $|B_n - B_n'| = o(|\phi'(b_n)|)$ $\forall v$ by Proposition 1. Thus the first term on the right-hand side of the inequality in (93) is $o(|\phi'(b_n)|)$, $\forall v$ and $\forall w$. For the second term in the inequality in (93), $|B_n'|$ is bounded $\forall v$ by Proposition 1 and $|A_n - A_n'| = o(r_n)$ by Feller (1971), Chapter XVI, Section 2, Theorem 1 uniformly in w. Thus this second term is $o(r_n)$, $\forall v$ and $\forall w$. Thus the right-hand side of (93) is $o(\max\{r_n, |\phi'(b_n)|\})$, $\forall v$ and $\forall w$.

The result also holds when the $r_n(e^{-v} - 1)w$ term – now on the interval $|v| \leq e_n, \forall w$ – is included since

(a) $|wA_n|$ is bounded $\forall w$ by Theorem 17 of Petrov(1975).

(b) $|e^{-v} + 1|\,|B_n - B_n'| = o(|\phi'(b_n)|)$ by Corollary 2, note now on $|v| \leq -K \log|\phi'(b_n)|$.

(c) $|e^{-v} + 1|\,|B_n'|$ is bounded by Corollary 2, note now on $|v| \leq -K \log|\phi'(b_n)|$.

(d) $|w|\,|A_n - A_n'| = o(r_n)$ again $\forall w$ by Theorem 17 of Petrov(1975).

Hence, the result (64) for $|v| \leq -K \log|\phi'(b_n)|$ and $\forall w$.

For the case when $x_o = \infty$, the term $\frac{\kappa_3}{6\sigma^3\sqrt{n}}(w^3 - 3w)$ need not be included with the higher order terms. The reason is that this term is $o(r_n)$ by the definition of r_n when $x_o = \infty$, see (61). Therefore, (63) holds.

□

Proof of Corollary 7: By definition $H(v; 0, 1, k) = \exp[-(1 - kv)^{1/k}]$ so that

$$H'(v; 0, 1, k) = -(1 - kv)^{(1/k - 1)}\ \exp[-(1 - kv)^{1/k}].$$

Looking at $-\log H'$,

$$-\log H'(v) = \left(1 - \frac{1}{k}\right)\log(1 - kv) + (1 - kv)^{1/k}. \tag{94}$$

Expanding $(1-kv)^{1/k}$ by $e^{-v}(1-\frac{kv^2}{2})+O(k^2)$ and $(1-\frac{1}{k})\log(1-kv)$ by $v-kv+\frac{kv^2}{2}+O(k^2)$, the equation (94) is equal to

$$v+e^{-v}+k(\frac{v^2}{2}-v-\frac{e^{-v}v^2}{2})+O(k^2). \tag{95}$$

Looking at the log of equation (65),

$$-\log\left[\Lambda'(v)\{1+(\frac{v^2}{2}-v-\frac{v^2e^{-v}}{2})\phi'(b_n)\}\right]=v+e^{-v}-(\frac{v^2}{2}-v-\frac{e^{-v}v^2}{2})\phi'(b_n)+O(\phi(b_n)^2). \tag{96}$$

With the definition $k_n=-\phi'(b_n)$, (95) and (96) match up to $O(\phi'(b_n)^2)$. □

6 Simulation project

The main results of this paper have been concerned with limiting forms of the joint density of the sample sum and maximum, but there are many other quantities of interest in connection with this problem. In particular, we might be interested in the conditional mean, $E\{M_n|S_n\}$. In this section, we give both theoretical and simulation results for this.

The theoretical results are to some extent heuristic, since we do not have a rigorous proof that our asymptotics for joint densities extend to the calculation of conditional means. Nevertheless, it is natural to conjecture the result by simply integrating out the approximate conditional density of sample maximum given the sample sum that follows from our main results. Here we perform this calculation, and illustrate the result by simulation.

For the marginal distribution of sample maxima, it is natural to assume the penultimate approximation (Cohen 1982b, Gomes 1984), since many theoretical and practical results have shown that it is a better approximation in practice than the Gumbel approximation. Under the first-order approximation that the sample mean and the sample sum are independent, $E\{M_n|S_n\}$ is of course the same as the unconditional mean $E\{M_n\}$, and may be approximated by the mean of the approximating penultimate distribution. Here we improve this by calculating the $O(r_n)$ correction.

Proposition 7 implies the approximation

$$f_{S_n^*,M_n^*}(w,v)=\mathcal{N}'(w)(1-k_nv)^{1/k_n-1}\exp\{-(1-k_nv)^{1/k_n}\}[1-r_n(e^{-v}-1)w+o\{\max(r_n,|\phi'(b_n)|)\}]$$

where $k_n=-\phi'(b_n)$. Since $f_{S_n^*}(w)\approx\mathcal{N}'(w)$ with error $O(n^{-1/2})$, this leads to

$$f_{M_n^*|S_n^*}(v|w)=(1-k_nv)^{1/k_n-1}\exp\{-(1-k_nv)^{1/k_n}\}[1-r_n(e^{-v}-1)w+o\{\max(r_n,|\phi'(b_n)|)\}]. \tag{97}$$

Formally integrating (97),

$$E\{M_n^*|S_n^*=w\}=\int v(1-k_nv)^{1/k_n-1}\exp\{-(1-k_nv)^{1/k_n}\}[1-r_n(e^{-v}-1)w+o\{\max(r_n,|\phi'(b_n)|)\}]dv \tag{98}$$

where the integral is over $\{v:\ 1-k_nv>0\}$.

Unfortunately (98) is not analytically integrable as it stands. However, we may replace e^{-v} by $(1 - k_n v)^{1/k_n}$ with error of $O(k_n)$, so (98) may be rewritten

$$E\{M_n^*|S_n^* = w\} = \int v(1-k_n v)^{1/k_n - 1} \exp\{-(1-k_n v)^{1/k_n}\}[1 - r_n\{(1-k_n v)^{1/k_n} - 1\}w dv] + o[\max\{r_n, |\phi'(b_n)|\}] \quad (99)$$

where we are assuming (without rigorous proof) that the interchange of limit and integration in (99) is valid.

The integral in (99) may be evaluated analytically as

$$E\{M_n^*|S_n^* = w\} \approx \frac{\{1 - \Gamma(1+k_n)\}}{k_n} + r_n w \Gamma(1+k_n)$$

or equivalently

$$E\{M_n|S_n\} \approx b_n + a_n \frac{\{1 - \Gamma(1+k_n)\}}{k_n} + a_n r_n \frac{S_n - n\mu}{\sqrt{n\sigma^2}} \Gamma(1+k_n). \quad (100)$$

If we ignore the higher-order term, then we derive the simpler approximation

$$E\{M_n|S_n\} \approx b_n + a_n \frac{\{1 - \Gamma(1+k_n)\}}{k_n}. \quad (101)$$

We fit (100) and (101) to the following simulated data. We simulated 10,000 samples of size $n = 30, 75, 120, 360$ from two distributions in the domain of attraction of the Gumbel domain which have at least a twice (here, infinitely) differentiable densities: the standard Normal($\mu = 0, \sigma^2 = 1$) and Lognormal with mean, $\mu = e^{1/8}$, and variance, $\sigma^2 = e^{1/4}(e^{1/4} - 1)$. In each graph, we plot sum versus maximum. We draw a smooth curve through the scatterplot using the S-PLUS 'loess' function. We include 95% confidence bands to the data. Thus the "data" line estimates the "average" maximum along the possible sum values. We substituted in the following variables into (100) and (101):

Dist	Density	Inverse	b_n	a_n	k_n	r_n
Normal $\mathcal{N}$	$\mathcal{N}'$	$\mathcal{N}^{-1}$	$\mathcal{N}^{-1}(1 - 1/n)$	$\frac{1}{n \times \mathcal{N}'(b_n)}$	$1 - a_n \times b_n$	$\frac{b_n}{\sqrt{n}}$
Lognorm $\mathcal{L}$	$\mathcal{L}'$	$\mathcal{L}^{-1}$	$\mathcal{L}^{-1}(1 - 1/n)$	$\frac{1}{n \times \mathcal{L}'(b_n)}$	$1 - \frac{a_n}{b_n} \times \{1 + \frac{\log(b_n) - \mu}{(\sigma)^2}\}$	$\frac{b_n}{\sqrt{n\sigma^2}}$

Note that the calculations for b_n in the simulation are based on $1 - F$ as opposed to the theoretical result which is based on its asymptotic equivalence $-\log F$. The difference between $1 - F$ and $-\log F$ is smaller than the higher order term that is being studied.

From the graphs (see Figure 1), first we see that the approximation performs better in the lognormal case than in the normal case. However, this is to be expected because papers about penultimate approximations show that the rate of convergence is better for the lognormal distribution than for the normal distribution. Second even when $n = 30$, the fit for higher order expansion term (the linear fit) is a substantial improvement over the asymptotic independence result (the constant). Nevertheless, we note that the approximation does not lie within the 95% confidence bounds. Third, in the neighborhood of $n = 75$, the higher order term expansion provides a good fit to the data for

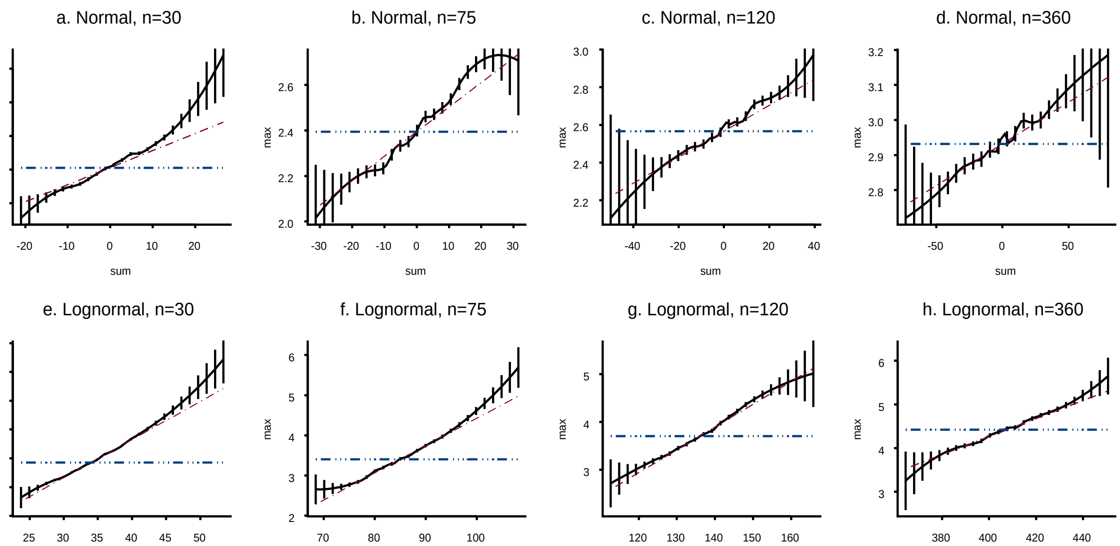

Figure 1: Simulation plots of sum versus maximum: Solid line is simulated data (loess fit with 95% confidence bands); Long dashed straight line is asymptotic Gumbel expected value $E[M_n]$, approximated in (101); Dotted line is expected value of $E[M_n|S_n]$ based on the approximation in (100) which uses the higher order term.

a substantial range of the data. The most prominent range of noncompliance is in the upper tail of the graph. This is more pronounced when $n = 120$. Fourth, even at a sample size equivalent to modelling "annual" data ($n = 360$) the asymptotic independence result still has not been realized. The Gumbel limit still does not lie within the confidence bands of the simulated data. Again, the higher order expansion does everywhere except the very upper tail.

A final comment relates to the systematic deviations with respect to the expansion's linear fit. In each case, the expansion fits better (lies more within the confidence bands) in the lower part of the graphs. In other words, the relationship between the sum and the maximum is approximately linear in the lower two-thirds of the graph but not in the upper tail. An interpretation of this is that the maxima are more compact for smaller sums. In other words, when the sum increases in the upper tail the maximum increases more rapidly than a linear relationship, implying for larger sums there is a higher probability of getting a very large maximum.

In conclusion, even at the annual sample level ($n = 360$) the asymptotic independence is not realized; that is, for moderate to moderately large samples the sum and the maximum are not independent. In fact, the simulation illustrates that the higher order expansion not only provides a substantial improvement over the asymptotic result but also that the fit for the expansion is within the 95% confidence bands for a large range of the data for moderate sample size. Thus the higher order expansion should provide a better model not only with the conditional case but in modelling the joint density as well. Finally, in the upper tails we have seen that the relationship between the sum and the maximum is not well approximated with this linear fit. An interesting extension to this work would be calculating a second higher order term to study if this second term would compensate for this nonlinear behavior in the upper tail.

7 Concluding Remarks

For moderate sample sizes, the higher order expansion term for the joint density of the sum and the maximum appears to model this dependence between the sum and maximum well for a large range of values. This should be a beneficial tool for modelers who either need to model this dependence or who are trying to model the maximum and have information on the sum. With respect to the upper tails where this linear approximation between the sum and the maximum does not appear to be a good fit, further research should include investigating how much benefit would there be in deriving the second higher order term in modelling the dependence between the sum and the maximum.

The present paper has not discussed possible applications of the results, but there are a number of contexts in which the joint distribution of sums and maxima could be useful in statistics. One particular context is climatology, where there is by now much research on trends in both the mean and the extremes of temperature and precipitation series. It is natural to model the means and extremes jointly, in order to test hypotheses such as whether there is a common trend that applies to both the means and the extremes. However, assuming that the annual mean and the annual maximum of a climatological series are statistically independent does not seem realistic for this kind of application. In preliminary work, approximations based on the results of this paper have performed much better in statistical work. Full details will be presented in a future publication.

Appendix

A Lemmas for Proposition 1

Lemma 8 *Let F be in the domain of attraction of Λ so that the representation in (3) holds with $c(x)=1$. Recall a_n and b_n are the appropriate normalizing constants for the Gumbel distribution defined in (5) and (4). Here $u_n = a_n v + b_n$. Also assume that*

$$\frac{\phi'(a_n v + b_n)}{\phi'(b_n)} \to 1, \quad \textit{uniformly on } |v| \le e_n = -K \log|\phi'(b_n)| \tag{102}$$

then

$$\frac{-\log F(u_n)}{-\log F(b_n)} = e^{-v}[1 + O(v^2|\phi'(b_n)|], \quad \textit{uniformly on } |v| \le e_n. \tag{103}$$

Remark The conditions of Proposition 1 are stronger and thus the results for Lemma 8 follow immediately under Proposition 1.

Proof: Here the argument is similar to that in the proof for Proposition 9.2 of Smith (1987).

From (3),

$$\frac{-\log F(u_n)}{-\log F(b_n)} = \exp\left\{-\int_0^v \frac{\phi(b_n)}{\phi(a_n t + b_n)}dt\right\}. \tag{104}$$

A Taylor expansion for the denominator of the integrand gives

$$\int_0^v \frac{\phi(b_n)}{\phi(a_n t + b_n)}dt = \int_0^v \frac{\phi(b_n)}{\phi(b_n) + a_n t\theta}dt$$

where $\theta = \phi'(a_n s + b_n)$ for some s between 0 and v. Recall that $\phi(b_n) = a_n$. Thus (104) is equal to

$$\exp\{-\frac{1}{\theta}\log(1+\theta v)\}. \tag{105}$$

Using a Taylor's expansion of $\log(1+x)$, (105) is equal to

$$\exp\{-(v + O(v^2\theta)\}. \tag{106}$$

Equation (102) allows the substitution of $\phi'(b_n)$ for θ in (106). Now using $e^{-v+r} = e^{-v}e^r$ and a Taylor expansion for $e^r = 1 + r + o(r)$ when $r \to 0$,

$$\frac{-\log F(u_n)}{-\log F(b_n)} = e^{-v}[1 + O(v^2|\phi'(b_n)|)], \quad \text{uniformly on } |v| \leq e_n.$$

□

Lemma 9 *Let $m \geq 1$ be a finite constant, $\delta > 0$ be an arbitrary finite constant, and κ be a finite constant. The function $h(v) = e^{-e^{-v}}e^{-mv}$ is uniformly bounded $\forall v$. In fact,*

$$h(v) \leq \min(e^{-m}m^m, \kappa|v|^{-\delta}), \quad \forall v. \tag{107}$$

Proof: The result falls from standard calculus calculations. Since the function $h(v)$ is a continuous on $-\infty < v < \infty$, it is straightforward to show that $h(v)$ has a finite sup and finite inflection points for any finite m. To resolve the exact form of the bound, divide the domain of h into two parts: $|v| \leq$ its inflection points and than $|v| >$ the inflection points. On the first interval, the maximum is $e^{-m}m^m$. On the other two ends, it is possible to bound the tails by a polynomial tail.

B Lemmas for Proposition 3

Lemma 10 *Let f be any integrable density, then $|\varphi_{u_n}(t)| \to 0$ uniformly in n as $|t| \to \infty$.*

Remark: The difference between Lemma 10 and Feller(1971), Section XV.4, Lemma 3 is that here the characteristic function is based on the conditional density.

Proof: In the conditional version,

$$|\varphi_{u_n}(t)| < \left| \frac{1}{F(u_n)} \int_{x_l}^{u_n} \cos(tx) f(x) dx \right| + \left| \frac{1}{F(u_n)} \int_{x_l}^{u_n} \sin(tx) f(x) dx \right|. \tag{108}$$

The modifications due to the conditional density in equation (108) is the $\frac{1}{F(u_n)}$ term and the upper limit of integration. Since u_n is a threshold (i.e. $u_n > x_l + \delta, \forall n$ and for some $\delta > 0$), $sup_n \frac{1}{F(u_n)} < \infty$. Hence $\frac{1}{F(u_n)}$ does not impact the uniformity results. As for the limits of integration, only positive functions are integrated in the proof. Thus if the result holds uniformly for the entire domain as in Feller (1971), then it holds when the integration extends just to u_n.

□

Lemma 11 *Let f' exist and be integrable, then*

$$\limsup_{n \to \infty} \sup_t |t| \; |\varphi_{u_n}(t)| < \infty.$$

Proof: This is an extension of Lemma 10. Again we modify the unconditional version – in this case, Feller (1971), Section XV.5, Lemma 4. The modifications include the $\frac{1}{F(u_n)}$ term and the upper limit of integration which we have seen from Lemma 10 does not affect the results. Also, we need the result only for $\limsup_{n \to \infty}$, a weaker result.

□

Corollary 12 *Given*

$$\limsup_{n \to \infty} \sup_t |t| \; |\varphi_{u_n}(t)| < \infty \tag{109}$$

then there exists an n^* such that $|\varphi_{u_n}(t)|^n$ is integrable for $n \geq n^* > 1$.

Proof: Clearly if (109) holds, then there exists an n^* and a constant $c < \infty$ such that for all $n \geq n^*$,

$$\sup_t |t| \; |\varphi_{u_n}(t)| \leq c < \infty.$$

This implies $\forall n \geq n^*$

$$|\varphi_{u_n}(t)| \leq \begin{cases} \frac{c}{|t|} & \text{if } |t| \geq c, \\ 1 & \text{if } |t| \leq c. \end{cases}$$

Note the last line follows since $\varphi_{u_n}(t)$ is a characteristic function. Therefore

$$\int_{-\infty}^{\infty} |\varphi_{u_n}(t)|^n dt \leq 2c + 2c^n \int_c^{\infty} \frac{dt}{t^n} = 2c + \frac{2c}{n-1} < \infty \qquad \forall n \geq n^*.$$

□

Corollary 13 *If $|\varphi_{u_n}(t)|^n$ is integrable for some $n \geq n^* > 1$, then $f_{\tilde{S_n}}$ exists and $\forall n \geq n^* > 1$ has Fourier norm*

$$N_n = \frac{1}{2\pi} \int_{-\infty}^{\infty} \left| \varphi_{u_n}^{n-1}\left(\frac{t}{\sqrt{(n-1)\sigma^2(u_n)}}\right) \exp\left(-\frac{it\mu(u_n)\sqrt{n-1}}{\sigma(u_n)}\right) \right| dt.$$

where n^* is defined in Corollary 12.

Proof: This falls from an application of the Fourier inversion formula of Feller (1971), Chapter XV, Section 3, Theorem 3.

The characteristic function of $\tilde{S_n}$ is

$$\begin{aligned} \varphi_{\tilde{S_n}}(t) &= E \exp\left(it \left[\frac{\sum_{i=1}^{n-1} X_i^* - (n-1)\mu(u_n)}{\sqrt{(n-1)\sigma^2(u_n)}} \right] \right) \\ &= \exp\left(-\frac{it\mu(u_n)\sqrt{n-1}}{\sigma(u_n)}\right) \; \varphi_{u_n}\left(\frac{t}{\sqrt{(n-1)\sigma^2(u_n)}}\right)^{n-1} \end{aligned}$$

where X^* has the same distribution $X|X \leq u_n$.
Corollary 12 shows that this is integrable thus the Fourier inversion theorem holds. □

Lemma 14 *For a continuous underlying distribution, given a $\delta > 0$ there exists a number $q_\delta < 1$ such that*

$$|\varphi_{u_n}(t)| < q_\delta \quad \forall |t| > \delta \text{ and } \forall n \geq n^*. \tag{110}$$

Proof: Using the same arguments as in the proofs of Lemma 10 and 11 to deal with $\frac{1}{F(u_n)}$ and the limits of integration in the definition of $\varphi_{u_n}(t)$, it is straightforward to show

$$\lim_{n\to\infty} |\varphi_{u_n}(t) - \varphi(t)| = 0, \text{uniformly in } t. \tag{111}$$

By standard properties of characteristic functions, see Feller(1971), Chapter XV, Section 1,

$$|\varphi(t)| < 1 \quad \text{whenever} \quad |t| \neq 0 \tag{112}$$

and

$$\varphi(t) \ \text{ is continuous.} \tag{113}$$

From the Riemann-Lebesgue Theorem, see Feller (1971), Chapter XV, Section 4, Lemma 3,

$$\exists\ \delta^* \ \text{ such that } \ |\varphi(t)| < \frac{1}{2} \ \ \forall |t| > \delta^*. \tag{114}$$

Now fix $\delta > 0$. Using (112) and (113),

$$\exists\ q_{\delta^*} \in (\frac{1}{2}, 1) \text{ such that } \ |\varphi(t)| < q_{\delta^*} \ \text{ on } \ \delta \le |t| \le \delta^*. \tag{115}$$

Using (114) and (115),

$$|\varphi(t)| < q_\delta^* \ \ |t| \ge \delta.$$

Now, let $q_\delta = (q_\delta^* + 1)/2$. Then by (111), choose n^* sufficiently large so that

$$|\varphi_{u_n}(t) - \varphi(t)| \le q_\delta - q_\delta^* \quad \forall t, \quad \forall n \ge n^*.$$

Using a triangular inequality argument,

$$|\varphi_{u_n}(t)| < q_\delta \ \ \forall |t| \ge \delta, \quad \forall n \ge n^*.$$

□

Recall the notation

$$\psi(t) = \log\varphi(t) - it\mu + \frac{t^2}{2}\sigma^2, \qquad \psi_{u_n}(t) = \log\varphi_{u_n}(t) - it\mu(u_n) + \frac{t^2}{2}\sigma^2(u_n)$$

$$\mathcal{K}_3 = \mu_3 - 3E(X^2)\mu + 2\mu^3, \qquad \mathcal{K}_3(u_n) = \mu_3(u_n) - 3E(X^2|X < u_n)\mu(u_n) + 2\mu(u_n)^3$$

Lemma 15 *Assume μ_3 exists, there exists a $\delta > 0$ such that φ''' exists and is continuous for some $|t| < \delta$, and $|\varphi_{u_n}(t)|^{n^*}$ is integrable for some $n^* > 1$, then*

$$|\psi_{u_n}(t) - \frac{(it)^3\mathcal{K}_3(u_n)}{6}| < \epsilon|t|^3, \quad \forall |t| < \delta, \forall n \ge n^*. \tag{116}$$

Proof: The inequality in equation (116) is solved by looking at the three term Taylor expansion of ψ_{u_n} and putting a bound on the remainder term. To do so, identifying the coefficients in the Taylor's expansion involves the following moments and characteristic functions.

Given the existence of μ_3, clearly the lower moments and the corresponding conditional moments exist. In fact, by dominated convergence theorem,

$$\mu_3(u_n) \to \mu_3, \quad E(X^2|X < u_n) \to E(X^2), \quad \mu(u_n) \to \mu \text{ and } \mathcal{K}_3(u_n) \to \mathcal{K}_3.$$

Also by a similar argument since φ''' exists and is continuous, then φ'' and φ' exist and are continuous as are the corresponding conditional versions. In fact,

$$\varphi_{u_n}{}''' \to \varphi''', \quad \varphi_{u_n}{}'' \to \varphi'', \text{and } \varphi_{u_n}{}' \to \varphi'.$$

By usual characteristic function properties,

$$\varphi_{u_n}(0) = 1, \quad \varphi_{u_n}{}'(0) = i\mu(u_n), \ \varphi_{u_n}{}''(0) = i^2 E(X^2|X < u_n), \ \varphi_{u_n}{}'''(0) = i^3 E(X^3|X < u_n).$$

Using the definition of ψ_{u_n}

$$\psi_{u_n}{}'''(t) = \frac{\varphi_{u_n}{}'''(t)}{\varphi_{u_n}(t)} - 3\frac{\varphi_{u_n}{}''(t)\varphi_{u_n}{}'(t)}{\varphi_{u_n}{}^2(t)} + 2(\frac{\varphi_{u_n}{}'(t)}{\varphi_{u_n}(t)})^3. \tag{117}$$

When $t = 0$,

$$\psi_{u_n}(0) = \psi_{u_n}{}'(0) = \psi_{u_n}{}''(0) = 0 \text{ with } \psi_{u_n}{}'''(0) = i^3 \mathcal{K}_3(u_n). \tag{118}$$

Taking the Cauchy form of Taylor's expansion (see Johnson and Kotz (1982), Vol. 9, p. 187) of $\psi_{u_n}(t)$ about 0 and substituting in the coefficients found in (118),

$$\left|\psi_{u_n}(t) - \frac{(it)^3\mathcal{K}_3(u_n)}{6}\right| = \left|\frac{t^3}{6}\{\psi_{u_n}{}'''(\theta) - \psi_{u_n}{}'''(0)\}\right| \tag{119}$$

for some $\theta \in (0, 1)$.

To put the appropriate bound on the RHS of equation (119), the form of ψ_{u_n} in (117) yields that it is sufficient to show for each $m = 0, 1, 2, 3$ that given $\epsilon > 0$ there exists a $\delta > 0$ (and $n^* > 1$) such that

$$|\varphi_{u_n}{}^{(m)}(t) - \varphi_{u_n}{}^{(m)}(0)| < \epsilon \quad \forall |t| < \delta \text{ and } \forall n \geq n^*. \tag{120}$$

In fact, the LHS of (120) can be divided into the following three parts:

$$|\varphi_{u_n}{}^{(m)}(t) - \varphi_{u_n}{}^{(m)}(0)| \leq |\varphi_{u_n}{}^{(m)}(t) - \varphi^{(m)}(t)| + |\varphi^{(m)}(t) - \varphi^{(m)}(0)| + |\varphi_{u_n}{}^{(m)}(0) - \varphi^{(m)}(0)|. \tag{121}$$

For the first term in the RHS of (121),

$$\left|\varphi_{u_n}{}^{(m)}(t) - \varphi^{(m)}(t)\right| \leq \left|\varphi^{(m)}(t)\right| \left(\frac{1-F(u_n)}{F(u_n)}\right) + \frac{1}{F(u_n)} \left|\int_{u_n}^{\infty} (ix)^m e^{itx} f(x) dx\right|. \quad (122)$$

Clearly the first term of the RHS of (122) tends to 0 as $n \to \infty$ since $\frac{1-F(u_n)}{F(u_n)} \to 0$ and $|\varphi^{(m)}(t)|$ is bounded for $m = 0, 1, 2, 3$. The second term also tends to 0, by dominated convergence theorem. Thus there exists a n^* such that $\forall n \geq n^*$

$$|\varphi_{u_n}{}^{(m)}(t) - \varphi^{(m)}(t)| < \epsilon/3 \quad \forall t. \quad (123)$$

For the second term in (121), due to the continuity of $\varphi^{(m)}$ in a neighborhood of 0, for each m, given an $\epsilon > 0$ there exists a $\delta > 0$ such that

$$|\varphi^{(m)}(t) - \varphi^{(m)}(0)| < \epsilon/3 \ \ \forall |t| < \delta. \quad (124)$$

Finally, for the third term in (121), recall $\varphi_{u_n}{}^{(m)}(0) \to \varphi^{(m)}(0)$ as $n \to \infty$ by dominated convergence theorem. Thus for each m, there exists a n^* such that

$$|\varphi_{u_n}{}^{(m)}(0) - \varphi^{(m)}(0)| < \epsilon/3 \ \ \forall n \geq n^* \ \ \forall t. \quad (125)$$

Putting (123), (124), and (125) together, define δ as smallest necessary in (124) and n^* as large as necessary in (123) and (125). Then (121) holds. □

C Lemmas for Proposition 4

Lemma 16 *Let F be in the domain of attraction of Λ so that representation in (3) holds with $c(x) = 1$, then as $u \to x_o$*

$$\phi(u)/u \to 0, \quad x_o = \infty \quad (126)$$

or

$$\phi(u) \to 0, \quad x_o < \infty. \quad (127)$$

Proof: If $x_o < \infty$, then $-\log F(x_o) = 0$ so in the Balkema and de Haan (1972) representation [(3) with $c(x) = 1$],

$$\exp\left\{-\int_0^{x_o} \frac{dt}{\phi(t)}\right\} = 0.$$

Assume $\phi(t) \not\to 0$ as $t \to x_o$. Since ϕ is continuous, if $\phi(x_o) \neq 0$, then there exists a constant $c \neq 0$ such that given an $\epsilon > 0$ there exists a t_c such that for some $t > t_c$

$$c - \epsilon < |\phi(t)| < c + \epsilon.$$

The implication is

$$\exp\left\{-\int_0^{x_o} \frac{dt}{\phi(t)}\right\} > \exp\left\{-\int_0^{t_c} \frac{dt}{\phi(t)}\right\} + \exp\left\{-\int_{t_c}^{x_o} \frac{dt}{c-\epsilon}\right\} > \exp\left\{-\frac{x_o - t_c}{c - \epsilon}\right\}.$$

This last term is a constant greater than 0 and thus a contradiction. Therefore

$$\phi(u) \to 0, \quad x_o < \infty.$$

For the case when $x_o = \infty$: The definition of the Balkema and de Haan (1972) representation (3) with $c(x) = 1$ includes $\phi'(u) \to 0$ as $u \to \infty$. This implies that given a $\delta > 0$ there exists a u_δ such that

$$|\phi'(u)| < \frac{\delta}{2}, \ \forall u > u_\delta. \tag{128}$$

Now assume $\frac{\phi(u)}{u} \not\to 0$. This implies that exists an infinite sequence of u such that along this sequence

$$\phi(u) > \kappa_1 u \quad \text{for some } \kappa_1 > 0. \tag{129}$$

By the Fundamental Theorem of Calculus and using (128)

$$\begin{aligned} \phi(u) &= \phi(u_\delta) + \int_{u_\delta}^{u} \phi'(t)dt \\ &< \phi(u_\delta) + \frac{\delta}{2}\{u - u_\delta\} \\ &\leq \kappa_2 + \frac{\delta}{2}u. \end{aligned} \tag{130}$$

In the last line we are using that both u_δ and $\phi(u_\delta)$ are some finite constants.

Combining (129) and (130),

$$u < \frac{\kappa_2}{\kappa_1 - \frac{\delta}{2}} = \kappa_3$$

where κ_3 is some finite constant. But $u \to x_o = \infty$. Given this contradiction, the conclusion is

$$\frac{\phi(u)}{u} \to 0, \quad x_o = \infty.$$

□

Bibliography

Anderson, C.W. (1971). *Contributions to the Asymptotic Theory of Extreme Values.* Ph.D. Thesis, University of London.

Anderson, C.W. and Turkman, K.F. (1991). The joint limiting distribution of sums and maxima of stationary sequences. *J. Appl. Prob.* **28**, 33-44.

Balkema, A.A. and de Haan, L. (1972). On R. von Mises' condition for the domain of attraction of $\exp(-e^{-x})$. *Ann. Math. Statist.* **43** 1352-1354.

Chow, T.L. and Teugels, J.L. (1978). The sum and the maximum of i.i.d. random variables. In *Proc. Second Prague Symp. Asymptotic Statistics*, 81-92.

Cohen, J.P. (1982a). The penultimate form of approximation to normal extremes. *Adv. Appl. Probab.* **14** 324-339.

Cohen, J.P. (1982b). Convergence rates for the ultimate and penultimate approximations in extreme-value theory. *Adv. Appl. Probab.* **14** 833-854.

Feller, W. (1971). *An Introduction to Probability Theory and Its Applications.* Vol. II, Wiley, New York.

Gomes, M.I. (1984). Penultimate limiting forms in extreme value theory. *Ann. Inst. Statist. Math.* **36** 71-85.

Grady, A. and Smith, R. L. (2003). A Higher Order Expansion for the Joint Density of the Sum and the Maximum under the Fréchet and Weibull domains of attraction. In preparation for submission.

Hall, P. (1979). On the rate of convergence of normal extremes. *J. Appl. Prob.* **16** 433-439.

Hall, P. (1980). Estimating probabilities for normal extremes. *Adv. Appl. Probab.* **12** 491-500.

Hall, P. (1982). *Rates of Convergence in the Central Limit Theorem.* Pitman Publishing, Inc., London.

Hall, W.J. and Wellner, J.A. (1979). The rate of convergence in law of the maximum of an exponential sample. *Statistica Neerlandica* **33** 151-154.

Johnson, N.L. and Kotz, S. (1982). *Encyclopedia of Statistical Sciences.* John Wiley, New York.

Leadbetter, M.R., Lindgren, G. and Rootzén, H. (1983). *Extreme and Related Properties of Random Sequences and Processes.* Springer-Verlag, New York.

Petrov, V.V. (1975). *Sums of Independent Random Variables.* Springer-Verlag, Berlin.

Pickands, J. (1975). Statistical inference using extreme order statistics. *Ann. Statist.* **3** 119-131.

Pickands, J. (1986). The continuous and differentiable domains of attraction of the extreme value distributions. *Ann. Probab.* **14** 996-1004.

Smith, R.L. (1987). Estimating tails of probability distributions. *Ann. Statist.* **15** 1174-1207.

Sweeting, T.J. (1985). On domains of uniform local attraction in extreme value theory. *Ann. Probab.* **13** 196-205.