Contribution of the Extreme Term in the Sum of Samples with Regularly Varying Tail

Van Minh Nguyen *

January 31, 2018

Abstract

For a sequence of random variables \((X_1, X_2, \ldots, X_n)\), \(n \geq 1\), that are independent and identically distributed with a regularly varying tail with index \(-\alpha, \alpha \geq 0\), we show that the contribution of the maximum term \(M_n \triangleq \max(X_1, \ldots, X_n)\) in the sum \(S_n \triangleq X_1 + \cdots + X_n\), as \(n \to \infty\), decreases monotonically with \(\alpha\) in stochastic ordering sense.

Keywords

Extreme · Sum · Regular variation · Stochastic ordering · Wireless network modeling

Mathematics Subject Classification (2000)

60G70 · 60G50 · 60E15 · 90B15

1 Introduction

Let \((X_1, X_2, \ldots)\) be a sequence of random variables that are independent and identically distributed (i.i.d.) with distribution \(F\). For \(n \geq 1\), the extreme term and the sum are defined as follows:

\[
M_n \triangleq \max_{i=1}^n X_i, \quad S_n \triangleq \sum_{n=1}^n X_i.
\]

The influence of the extreme term in the sum has various implications in both theory and applications. In particular, it has been used to characterize the nature of possible convergence of the sums of i.i.d. random variables (Darling 1952). On the other hand, besides well-known applications to risk management, insurance and finance (Embrechts et al. 1997), it has been recently applied to wireless communications for characterizing a fundamental parameter which is the signal-to-interference ratio (Nguyen et al. 2010; Nguyen & Kountouris 2017).

Denote by \(\bar{F}(x) \triangleq 1 - F(x)\) the tail distribution of \(F\), a primary result of this question is due to (Darling 1952) in which it is shown that

\*Huawei Technologies, France. Email: vanminh.nguyen@huawei.com
Suppose that \( X_i \geq 0 \). Then
\[
\mathbb{E}\left( \frac{S_n}{M_n} \right) \to 1 \quad \text{as} \quad n \to \infty
\]
if for every \( t > 0 \) we have
\[
\lim_{x \to \infty} \frac{\bar{F}(tx)}{\bar{F}(x)} = 1.
\]

In Darling’s result, condition (2) is of particular interest as it is a specific case of a more general class called regularly varying tails which is defined in the following (Bingham et al. 1989):

**Definition 1.** A positive, Lebesgue measurable function \( h \) on \((0, \infty)\) is called regularly varying with index \( \alpha \in \mathbb{R} \) at \( \infty \) if \( \lim_{x \to \infty} \frac{h(tx)}{h(x)} = t^\alpha \) for every constant \( 0 < t < \infty \).

In the sequel, \( R_\alpha \) denotes the class of regularly varying functions with index \( \alpha \), and in particular \( R_0 \) is referred to as the class of slowly varying functions. In addition, \( \xrightarrow{d}, \xrightarrow{p}, \) and \( \xrightarrow{a.s.} \) stands for the convergence in distribution, convergence in probability, and almost sure convergence, respectively.

Since the work of (Darling 1952), there has been subsequent extensions which in particular investigated other cases of regularly varying tails. Among those, (Arov & Bobrov 1960) derived the characteristic function and limits of the jointed sum and extreme term for a regularly varying tail. (Teugels 1981) derived the limiting characteristic function of the ratio of the sum to order statistics, and moreover investigated norming sequences for its convergence to a constant or a normal law. (Chow & Teugels 1978; Anderson & Turkman 1991, 1995) investigated the asymptotic independence of normed extreme and normed sum. Unlike the slowly varying case, (O’Brien 1980) showed that \( M_n/S_n \xrightarrow{a.s.} 0 \Leftrightarrow \mathbb{E}X_1 < \infty \). For \( \bar{F} \) regularly varying with index \(-\alpha, \alpha > 0\), (Bingham & Teugels 1981) showed that the extreme term only contributes a proportion to the sum:

**Property 1** (Darling 1952). Suppose that \( X_i \geq 0 \). Then \( \mathbb{E}(S_n/M_n) \to 1 \) as \( n \to \infty \) if for every \( t > 0 \) we have

\[
\lim_{x \to \infty} \frac{\bar{F}(tx)}{\bar{F}(x)} = 1.
\]

**Property 2** (Bingham & Teugels 1981). The following are equivalent:

1. \( \bar{F} \in R_{-\alpha} \) for \( 0 < \alpha < 1 \);
2. \( M_n/S_n \xrightarrow{d} R \) where \( R \) has a non-degenerate distribution;
3. \( \mathbb{E}(S_n/M_n) \to (1 - \alpha)^{-1} \).

**Property 3** (Bingham & Teugels 1981). Let \( \mu = \mathbb{E}X_1 \). The following are equivalent:

1. \( \bar{F} \in R_{-\alpha} \) for \( 1 < \alpha < 2 \);
2. \( (S_n - (n - 1)\mu)/M_n \xrightarrow{d} D \) where \( D \) has a non-degenerate distribution;
3. \( \mathbb{E}((S_n - (n - 1)\mu)/M_n) \to c \) where \( c \) is a constant.

(Maller & Resnick 1984) extended Darling’s convergence in mean to convergence in probability of \( S_n/M_n \):
Property 4 ((Maller & Resnick 1984)).

\[
\frac{S_n}{M_n} \overset{P}{\to} 1 \iff \bar{F} \in \mathcal{R}_0.
\]

The ratio of the extreme to the sum has been further studied with the following result:

Property 5 ((Downey & Wright 2007)). If either one of the following conditions:

1. \(\bar{F} \in \mathcal{R}_{-\alpha}\) for \(\alpha > 1\),
2. \(F\) has finite second moment,

holds, then

\[
\mathbb{E}(M_n/S_n) = \frac{EM_n}{ES_n}(1 + o(1)), \quad \text{as} \ n \to \infty.
\]

To this end, the contribution of the extreme term in the sum has been investigated and classified in the following cases:

- \(\bar{F} \in \mathcal{R}_0\);
- \(\bar{F} \in \mathcal{R}_{-\alpha}\), \(0 < \alpha < 1\);
- \(\bar{F} \in \mathcal{R}_{-\alpha}\), \(1 < \alpha < 2\);
- \(\bar{F} \in \mathcal{R}_{-\alpha}\), \(\alpha > 2\).

Nevertheless, how the influence of the extreme term in the sum gradually varies with the regular variation index has not been quantified and remains an open question. Precisely, consider two cases with \(\bar{F} \in \mathcal{R}_{-\alpha_1}\) and \(\bar{F} \in \mathcal{R}_{-\alpha_2}\) in which \(0 \leq \alpha_1 < \alpha_2\), which case results in larger \(M_n/S_n\)? This question is particularly important for analysis and design of wireless communication networks.

In this context, random variables \((X_1, X_2, \ldots)\) are used to model the signal that a user receives from base stations. It has been proven that the tail distribution of \(X_i\) is regularly varying (Nguyen & Kountouris 2017) either due to the effect of distance-dependent propagation loss or due to fading (Tse & Viswanath 2005) that is regular varying (Rajan et al. 2017) as a consequence of advances in communication and signal processing techniques such as massive multiple-input multiple-output transmission, coordinated multipoint and millimeter wave systems. Meanwhile, \(M_n\) expresses the useful signal power and \((S_n - M_n)\) is the total interference due to the other transmitters (Nguyen 2011). \(M_n/(S_n - M_n)\) is hence the signal-to-interference ratio (SIR), and its limit as \(n \to \infty\) happens for a dense or ultra-dense network (Nguyen & Kountouris 2016; Nguyen 2017). Capacity of a communication channel is expressed in term of the well-known Shannon’s capacity limit of \(\log(1 + \text{SIR})\) (Shannon 1948; Cover & Thomas 2006) considering that thermal noise is negligible in comparison to the interference. Therefore, \(M_n/S_n\) (or \(S_n/M_n\)) is a fundamental parameter of wireless network engineering. From the perspective of capacity, a primary purpose...
is to design the network such that $X_i$ possesses properties that make $M_n/S_n$ as large as possible. In particular, how $M_n/S_n$ varies according to the tail of $X_i$ turns out to be a critical question.

In this paper, we establish a stochastic ordering for $S_n/M_n$ and show that between two cases with $\bar{F} \in \mathbb{R}_{-\alpha_1}$ and $\bar{F} \in \mathbb{R}_{-\alpha_2}$ in which $0 \leq \alpha_1 < \alpha_2$, the contribution of $M_n$ in $S_n$ as $n \to \infty$ is larger in the former than in the latter case in stochastic ordering sense.

2 Main Result

In the following, for $n \geq 1$ we define

$$R_n \triangleq S_n/M_n.$$  \hfill (3)

We also restrict our consideration to non-negative random variables, i.e., $X_i \geq 0$, and consider $\bar{F} \in \mathbb{R}_{-\alpha}$ with $\alpha \geq 0$. In the context where $\alpha$ is analyzed, a variable $v$ is written as $v_\alpha$, e.g., write $S_{\alpha,n}$, $M_{\alpha,n}$, and $R_{\alpha,n}$ for $S_n$, $M_n$, and $R_n$, respectively.

Lemma 1. For $s \in \mathbb{C}$ with $\Re(s) \geq 0$, define $\mathcal{L}_{R_n}(s) \triangleq \mathbb{E}(e^{-sR_n})$. If $\bar{F} \in \mathbb{R}_{-\alpha}$ with $\alpha \geq 0$ then:

$$\mathcal{L}_{R_n}(s) = \frac{e^{-s}}{1 + \phi_\alpha(s)}, \quad n \to \infty,$$  \hfill (4)

where

$$\phi_\alpha(s) \triangleq \alpha \int_0^1 \frac{(1 - e^{-st}) dt}{t^{1+\alpha}}. \hfill (5)

Proof. Let $G(x_1, \cdots, x_n)$ be the joint distribution of $(X_1, \cdots, X_n)$ given that $M_n \geq X_1$, it is given as follows:

$$G(dx_1, \cdots, dx_n) = \begin{cases} F(dx_1) \cdots F(dx_n) & \text{if } x_1 = \max_{i=1}^n x_i, \\ 0 & \text{otherwise} \end{cases}.$$  \hfill (6)

Since $(X_1, \cdots, X_n)$ are i.i.d., $M_n = X_1$ with probability $1/n$. Thus, the (unconditional) joint distribution of $(X_1, \cdots, X_n)$ is $nG(x_1, \cdots, x_n)$. Hence,

$$\mathcal{L}_{R_n}(s) = \mathbb{E}(e^{-sR_n}) = n \int \cdots \int e^{-s(x_1 + \cdots + x_n)/x_1} G(dx_1, \cdots, dx_n),$$

and using $G$ from (6), we obtain

$$\mathcal{L}_{R_n}(s) = n \int_0^\infty \left( \int_0^x \cdots \int_0^x e^{-sx} \prod_{i=2}^n e^{-sx_i/x_i} F(dx_i) \right) F(dx)$$

$$= ne^{-s} \int_0^\infty \left( \int_0^1 e^{-st}F(xdt) \right)^{n-1} F(dx)$$

$$= ne^{-s} \int_0^\infty (\phi(x))^{n-1} F(dx),$$  \hfill (7)
where
\[ \varphi(x) \triangleq \int_0^1 e^{-st} F(x) dt. \] (8)

Given that \( \Re(s) \geq 0 \), we can see that
\[ |\varphi(x)| \leq \int_0^1 |e^{-st} F(x) dt| \leq \int_0^1 F(x) dt = F(x) < 1, \quad \text{for } x \to \infty. \]

Hence,
\[ \int_0^T (\varphi(x))^{n-1} F(dx) \to 0, \quad \text{as } n \to \infty \text{ for } T < \infty. \]

As a result, we only need to consider the contribution of large \( x \) in (7) for \( \mathcal{L}_{R_n}(s) \). An integration by parts with \( e^{-st} \) and \( F(x) dt \) yields:
\[ \varphi(x) = 1 - e^{-s} \tilde{F}(x) - \int_0^1 se^{-st} \tilde{F}(x) dt \]
\[ = 1 - \tilde{F}(x) + \int_0^1 se^{-st} (\tilde{F}(x) - \tilde{F}(tx)) dt. \] (9)

For \( \tilde{F} \in \mathcal{R}_{-\alpha} \) with \( \alpha \geq 0 \), we can write
\[ \tilde{F}(tx) \sim t^{-\alpha} \tilde{F}(x), \quad \text{as } x \to \infty, \quad 0 < t < \infty. \]

Thus
\[ \int_0^1 se^{-st} (\tilde{F}(x) - \tilde{F}(tx)) dt \sim \tilde{F}(x) \int_0^1 se^{-st} (1 - t^{-\alpha}) dt, \quad \text{as } x \to \infty. \]

Using an integration by parts with \( (1 - t^{-\alpha}) \) and \( d(1 - e^{-st}) \) we obtain
\[ \tilde{F}(x) \int_0^1 se^{-st} (1 - t^{-\alpha}) dt = -\tilde{F}(x) \int_0^1 \alpha(1 - e^{-st}) \frac{dt}{t^{1+\alpha}}. \] (10)

Put
\[ \phi_\alpha(s) \triangleq \int_0^1 \alpha(1 - e^{-st}) \frac{dt}{t^{1+\alpha}}, \]
and substitute it back in (10) and (9), we obtain
\[ \varphi(x) = 1 - (1 + \phi_\alpha(s)) \tilde{F}(x), \quad \text{for } x \to \infty. \]

Substitute \( \varphi(x) \) back in the expression of \( \mathcal{L}_{R_n}(s) \) in (7), we obtain
\[ \mathcal{L}_{R_n}(s) \sim ne^{-s} \int_0^\infty (1 - (1 + \phi_\alpha(s)) \tilde{F}(x))^{n-1} F(dx), \quad \text{as } n \to \infty. \] (11)
Here, we resort to a change of variable with \( v = n \bar{F}(x) \) and obtain:

\[
\mathcal{L}_{R_n}(s) \sim e^{-s} \int_0^n \left(1 - \frac{v}{n}(1 + \phi_\alpha(s))\right)^{n-1} dv
\]

\[
\xrightarrow{(a)} e^{-s} \int_0^\infty e^{-v(1+\phi_\alpha(s))} dv = \frac{e^{-s}}{1+\phi_\alpha(s)}, \quad \text{as } n \to \infty
\]

where (a) is due to the formula \((1 + \frac{x}{n})^n \to e^x\) as \( n \to \infty \).

To present the main result, we use the following notation. For two random variables \( U \) and \( V \), \( U \) is said to be smaller than \( V \) in Laplace transform order, denoted by \( U \preceq_{Lt} V \), if and only if \( \mathcal{L}_U(s) = \mathbb{E}(e^{-sU}) \geq \mathbb{E}(e^{-sV}) = \mathcal{L}_V(s) \) for all positive real number \( s \).

**Theorem 1.** Let \( R_{\alpha_1,n} \) and \( R_{\alpha_2,n} \) be as defined in (3) for \( \bar{F} \in \mathcal{R}_{-\alpha_1} \) with \( \alpha_1 \geq 0 \) and for \( \bar{F} \in \mathcal{R}_{-\alpha_2} \) with \( \alpha_2 \geq 0 \), respectively. Then

\[
\alpha_1 \leq \alpha_2 \Rightarrow R_{\alpha_1,n} \preceq_{Lt} R_{\alpha_2,n}, \quad n \to \infty.
\]

**Proof.** The proof is direct from Lemma 1. By noting that \( \alpha/t^{1+\alpha} \) is increasing with respect to (w.r.t.) \( \alpha \geq 0 \) for \( t \in [0,1] \), we have \( \phi_\alpha(s) \) in (5) is increasing w.r.t. \( \alpha \geq 0 \). Thus, \( \mathcal{L}_{R_{\alpha_1,n}}(s) \) as \( n \to \infty \) and \( \Re(s) > 0 \) is decreasing w.r.t. \( \alpha \geq 0 \).

It follows that, for \( 0 \leq \alpha_1 \leq \alpha_2 \), we have \( \mathcal{L}_{R_{\alpha_1,n}}(s) \geq \mathcal{L}_{R_{\alpha_2,n}}(s) \) as \( n \to \infty \) for all \( s \) with \( \Re(s) > 0 \), thus \( R_{\alpha_1,n} \preceq_{Lt} R_{\alpha_2,n} \) as \( n \to \infty \).

Theorem 1 dictates that the more slowly \( \bar{F} \) decays at \( \infty \), i.e., smaller \( \alpha \), the smaller is the ratio of the sum to the extreme, thus the bigger is the contribution of the extreme term in the sum. This contribution of the extreme term to the sum increases to the ceiling limit 1 when \( \alpha \) gets close to 0 as we have known from (Darling 1952; Maller & Resnick 1984).

Since we have established the Laplace transform ordering for \( R_{\alpha_1,n} \), an immediate application is related to completely monotonic and Bernstein functions. Let us recall:

- **Completely monotonic functions:** A function \( g : (0, \infty) \to \mathbb{R}_+ \) is said to be completely monotonic if it possesses derivatives of all orders \( k \in \mathbb{N} \cup \{0\} \) which satisfy \((-1)^kg^{(k)}(x) \geq 0, \forall x \geq 0\), where the derivative of order \( k = 0 \) is defined as \( g(x) \) itself. We denote by \( \mathcal{CM} \) the class of completely monotonic functions.

- **Bernstein functions:** A function \( h : (0, \infty) \to \mathbb{R}_+ \) with \( dh(x)/dx \) being completely monotonic is called a Bernstein function. We denote by \( \mathcal{B} \) the class of Bernstein functions.

Note that a completely monotonic function is positive, decreasing and convex, whereas a Bernstein function is positive, increasing and concave.
It is well known that for all completely monotonic functions \( g \), we have that \( U \preceq_{\mathcal{L}} V \Leftrightarrow \mathbb{E}(g(U)) \geq \mathbb{E}(g(V)) \), whereas for all Bernstein functions \( h \), \( U \preceq_{\mathcal{L}} V \Leftrightarrow \mathbb{E}(h(U)) \leq \mathbb{E}(h(V)) \). Hence, we can have a direct corollary of Theorem 1 as follows.

**Corollary 1.** With the same notation and assumption of Theorem 1, if \( 0 \leq \alpha_1 \leq \alpha_2 \), then:

\[
\forall g \in \mathcal{CM} : \quad \mathbb{E}(g(R_{\alpha_1,n})) \geq \mathbb{E}(g(R_{\alpha_2,n})), \quad n \to \infty,
\]
\[
\forall h \in \mathcal{B} : \quad \mathbb{E}(h(R_{\alpha_1,n})) \leq \mathbb{E}(h(R_{\alpha_2,n})), \quad n \to \infty.
\]

Note that \( h(x) = 1, \forall x > 0 \), is a Bernstein function, whereas \( g(x) = 1/x, \forall x > 0 \), is a completely monotonic function. For two cases with \( \bar{F} \in \mathcal{R}_{-\alpha_1} \) and \( \bar{F} \in \mathcal{R}_{-\alpha_2} \) with \( 0 \leq \alpha_1 \leq \alpha_2 \), Corollary 1 gives:

\[
\mathbb{E}\left( \frac{S_{\alpha_1,n}}{M_{\alpha_1,n}} \right) \leq \mathbb{E}\left( \frac{S_{\alpha_2,n}}{M_{\alpha_2,n}} \right), \quad n \to \infty,
\]
\[
\mathbb{E}\left( \frac{M_{\alpha_1,n}}{S_{\alpha_1,n}} \right) \geq \mathbb{E}\left( \frac{M_{\alpha_2,n}}{S_{\alpha_2,n}} \right), \quad n \to \infty.
\]

**Application Example** We now can show an application of the results developed above to the context of wireless communication networks. The signal-to-interference ratio SIR as described in the Introduction can be expressed in terms of \( R_n \) as follows:

\[
\frac{1}{\text{SIR}} = \frac{S_n - M_n}{M_n} = R_n - 1 := Z_n.
\]  \hspace{1cm} (13)

Assume that \( \bar{F} \in \mathcal{R}_{-\alpha} \), \( \alpha \geq 0 \), the Laplace transform of \( Z_n \) can be directly obtained from that of \( R_n \) as given by Lemma 1 as follows:

\[
\mathcal{L}_{Z_{\alpha,n}}(s) = e^{s} \mathcal{L}_{R_{\alpha,n}}(s) = (1 + \phi_\alpha(s))^{-1}, \quad n \to \infty,
\]  \hspace{1cm} (14)

for all \( s \in \mathbb{C}, \Re s > 0 \). It is easy to see that \( \mathcal{L}_{Z_{\alpha,n}}(s) \) is also decreasing with respect to \( \alpha \geq 0 \) (see the proof of Theorem 1).

Now, consider two cases for the distribution of the signal \( X_i \) which are \( \bar{F} \in \mathcal{R}_{-\alpha_1} \) and \( \bar{F} \in \mathcal{R}_{-\alpha_2} \) with \( 0 \leq \alpha_1 \leq \alpha_2 \), we firstly have:

\[
Z_{\alpha_1,n} \preceq_{\mathcal{L}} Z_{\alpha_2,n}, \quad n \to \infty.
\]  \hspace{1cm} (15)

Then, by noting that functions \( 1/x \) and \( \log(1 + 1/x) \) with \( x > 0 \) both are completely monotonic, we immediately have for \( n \to \infty \):

\[
\mathbb{E}(\text{SIR}_{\alpha_1}) \geq \mathbb{E}(\text{SIR}_{\alpha_2}), \quad (16)
\]
\[
\mathbb{E}(\log(1 + \text{SIR}_{\alpha_1})) \geq \mathbb{E}(\log(1 + \text{SIR}_{\alpha_2})). \quad (17)
\]

This says that it is beneficial to the communication quality and capacity to design the network such that the signal received from a transmitting base station \( X_i \) admits a regularly varying tail with as small variation index as possible (i.e., as close to a slowly varying tail as possible).
References

Anderson, C. W., & Turkman, K. F. 1991. The joint limiting distribution of sums and maxima of stationary sequences. *J. Appl. Probab.*, **28**(1), 33–44.

Anderson, C. W., & Turkman, K. F. 1995. Sums and maxima of stationary sequences with heavy tailed distributions. *Sankhyā: The Indian Journal of Statistics, Series A*, **57**(1), 1–10.

Arov, D. Z., & Bobrov, A. A. 1960. The extreme terms of a sample and their role in the sum of independent variables. *Theory Probab. Appl.*, **5**(4), 377–396.

Bingham, N. H., & Teugels, J. L. 1989. *Regular Variation*. Cambridge, U.K.: Cambridge University Press.

Cover, T. M., & Thomas, J. A. 2006. *Elements of Information Theory*. 2nd edn. John Wiley & Sons.

Darling, D. A. 1952. The influence of the maximum term in the addition of independent random variables. *Trans. Amer. Math. Soc.*, **73**(1), 95–107.

Downey, P. J., & Wright, P. E. 2007. The ratio of the extreme to the sum in a random sequence. *Extremes*, **10**(4), 249–266.

Embrechts, P., Klüppelberg, C., & Mikosch, T. 1997. *Modelling Extremal Events*. Berlin, Germany: Springer-Verlag.

Maller, R. A., & Resnick, S. I. 1984. Limiting behaviour of sums and the term of maximum modulus. *Proc. London Math. Soc.*, **3**(3), 385–422.

Nguyen, V. M. 2017. *Network Densification*. IET. Chap. 1, pages 3–24.

Nguyen, V. M., & Kountouris, M. 2016 (May). Coverage and capacity scaling laws in downlink ultra-dense cellular networks. *Pages 1–7 of: IEEE Int. Conf. Commun. (ICC)*.

Nguyen, V. M., & Kountouris, M. 2017. Performance limits of network densification. *IEEE J. Sel. Areas Commun.*, **35**(6), 1294–1308.

Nguyen, V. M., Baccelli, F., Thomas, L., & Chen, C. S. 2010. Best signal quality in cellular networks: asymptotic properties and applications to mobility management in small cell networks. *EURASIP J. Wirel. Commun. Netw.*, **2010**(May), 690161.
O’Brien, G. L. 1980. A limit theorem for sample maxima and heavy branches in Galton-Watson trees. *J. Appl. Probab.*, **17**(2)(Jun.), 539–545.

Rajan, A., Tepedelenlioğlu, C., & Zeng, R. 2017. A unified approach for modeling fading channels using infinitely divisible distributions. *IEEE Trans. Veh. Technol.*, **66**(11), 10194–10206.

Shannon, C. E. 1948. A mathematical theory of communication. *Bell Syst. Tech. J.*, **27**(3), 379–423.

Teugels, J. L. 1981. Limit theorems on order statistics. *Ann. Probab.*, **9**(5), 868–880.

Tse, David, & Viswanath, Pramod. 2005. *Fundamentals of Wireless Communication*. 1st edn. New York, U.S.A.: Cambridge University Press.