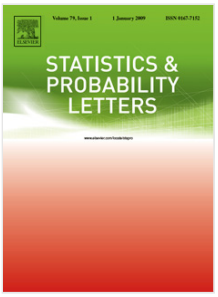
A uniform asymptotic expansion for weighted sums of exponentials

J.S.H. van Leeuwaarden, N.M. Temme

Accepted Manuscript



A uniform asymptotic expansion for weighted sums of exponentials

J.S.H. van Leeuwaarden, N.M. Temme

PII: S0167-7152(11)00188-X
DOI: 10.1016/j.spl.2011.05.013
Reference: STAPRO 6010

To appear in: *Statistics and Probability Letters*

Received date: 22 December 2010
Revised date: 23 May 2011
Accepted date: 25 May 2011

Please cite this article as: van Leeuwaarden, J.S.H., Temme, N.M., A uniform asymptotic expansion for weighted sums of exponentials. *Statistics and Probability Letters* (2011), doi:10.1016/j.spl.2011.05.013

A uniform asymptotic expansion for weighted sums of exponentials

J.S.H. van Leeuwaarden * N.M. Temme †

May 23, 2011

Abstract

We consider the random variable $Z_{n,\alpha} = Y_1 + 2^\alpha Y_2 + \ldots + n^\alpha Y_n$, with $\alpha \in \mathbb{R}$ and $Y_1, Y_2, \ldots$ independent and exponentially distributed random variables with mean one. The distribution function of $Z_{n,\alpha}$ is in terms of a series with alternating signs, causing great numerical difficulties. Using an extended version of the saddle point method, we derive a uniform asymptotic expansion for $\mathbb{P}(Z_{n,\alpha} < x)$ that remains valid inside ($\alpha \geq -1/2$) and outside ($\alpha < -1/2$) the domain of attraction of the central limit theorem. We discuss several special cases, including $\alpha = 1$, for which we sharpen some of the results in Kingman and Volkov (2003).

Key words: asymptotic analysis, saddle point approximation, exponential random variables, Gumbel distribution, Kolmogorov distribution, OK Corral model.

MSC2000 subject classification. 60F05, 33B20, 41A60.

1 Introduction

We consider the random variable

$$Z_{n,\alpha} = Y_1 + 2^\alpha Y_2 + \ldots + n^\alpha Y_n, \tag{1}$$

where $\alpha \in \mathbb{R}$ and $Y_1, Y_2, \ldots$ are independent and exponentially distributed random variables with mean one. It is fairly straightforward to derive an explicit expression for the distribution function $\mathbb{P}(Z_{n,\alpha} < x)$ in terms of an alternating series, see (4), but this series gives rise to severe numerical problems. In this paper we derive an expansion for the distribution function that will provide arbitrarily sharp approximations, uniformly for all $x \geq 0$. We derive the expansion by applying an extended version of the saddle point method. In order to do so, we first write $\mathbb{P}(Z_{n,\alpha} < x)$ as a contour integral that arises from applying the inversion formula to the Laplace transform of $Z_{n,\alpha}$. We then transform the integral and apply the saddle point method, in a way that is reminiscent of the saddle point approximation developed by Lugannani and Rice (1980) for the sum of i.i.d. random variables.

Apart from the case $\alpha = 0$, the sum in (1) consists of random variables having different distributions. Our primary motivation for undertaking this study is to demonstrate how the saddle

*Department of Mathematics and Computer Science, Eindhoven University of Technology, P.O. Box 513, 5600 MB Eindhoven, The Netherlands. E-mail: j.s.h.v.leeuwaarden@tue.nl

†CWI, Science Park 123, 1098 XG Amsterdam, The Netherlands. E-mail: nico.temme@cwi.nl

point method, along with some other asymptotic methods, leads to a tractable uniform expansion for $\mathbb{P}(Z_{n,\alpha} < x)$ that is sharp even for small values of n and x away from the mean. We have three more motivations for undertaking this effort:

- The random variable $Z_{n,\alpha}$ for the case $\alpha = 1$ occurs in various contexts, such as linear combinations of order statistics Holst (1980), Matsunawa (1985), noise in radio receivers Kac and Siegert (1947), and urn models Kingman (1999), Kingman and Volkov (2003). Building on the latter paper, Kingman and Volkov (2003) apply the central limit theorem (CLT) to $Z_{n,1}$ and derive the order of the error made by approximating a normalized version of $Z_{n,1}$ by a standard normal random variable. Our uniform expansion leads to arbitrarily sharp approximations.

- In fact, for all values $\alpha \geq -\frac{1}{2}$, the random variable $Z_{n,\alpha}$, after appropriate scaling, obeys the CLT and converges to a standard normal random variable. The resulting normal approximation to $\mathbb{P}(Z_{n,\alpha} < x)$ is useful for large n and x close to the mean of $Z_{n,\alpha}$. However, for increasing values of α, or for values of x far from the mean, the CLT kicks in ever more slowly. Our asymptotic expansion corrects for these effects.

- Our expansion remains valid for $\alpha < -1/2$. For $\alpha = -1$ the random variable $Z_{n,\alpha}$ describes the maximum of n i.i.d. exponentially distributed random variables. This is a classic example from extreme value theory, for which, as $n \to \infty$, $Z_{n,\alpha}$ follows the Gumbel distribution. For $\alpha = -2$, $Z_{n,\alpha}$ follows the Kolmogorov distribution, another classical distribution that plays for instance a role in the study of Brownian excursions in relation to the Riemann zeta function Biane et al. (2001).

In Section 2 we derive an integral representation for $\mathbb{P}(Z_{n,\alpha} < x)$ in terms of a contour integral in the complex plane. In Section 3, we use this integral representation to derive the asymptotic expansion for $\mathbb{P}(Z_{n,\alpha} < x)$, which holds uniformly for $x \geq 0$. In Section 4 we evaluate the coefficients in the expansion, in Section 5 we discuss several special cases, and in Section 6 we present some numerical results.

2 Integral representation

The Laplace transform $\nu_{n,\alpha}(s) = \mathbb{E}(e^{-sZ_{n,\alpha}})$ of $Z_{n,\alpha}$ can be represented as

$$\nu_{n,\alpha}(s) = \prod_{j=1}^{n} \frac{1}{1+sj^{\alpha}} = \sum_{j=1}^{n} \frac{1}{\Pi_{j,n,\alpha}} \frac{1}{1+sj^{\alpha}} \tag{2}$$

with

$$\Pi_{j,n,\alpha} = \prod_{1 \leq k \leq n, k \neq j} (1 - (k/j)^{\alpha}). \tag{3}$$

Let $F_{n,\alpha}(x) = \mathbb{P}(Z_{n,\alpha} > x)$. We show below that (2) leads to the alternative expression

$$F_{n,\alpha}(x) = \sum_{j=1}^{n} \frac{1}{\Pi_{j,n,\alpha}} e^{-xj^{-\alpha}}. \tag{4}$$

For the case $\alpha = 1$, we can write the distribution function as

$$F_{n,1}(x) = \frac{(-1)^n}{n!} \sum_{j=1}^{n} \binom{n}{j} (-1)^j j^n e^{-x/j}, \tag{5}$$

The limiting values at $x = 0$ and $x = \infty$ are given by

$$F_{n,1}(0) = 1, \quad F_{n,1}(\infty) = 0. \tag{6}$$

The value $F_{n,1}(\infty)$ follows easily from (5) and $F_{n,1}(0)$ follows from a relation with the Stirling numbers of the second kind, viz.

$$\mathcal{S}_n^m = \frac{1}{m!} \sum_{k=0}^{m} (-1)^{m-k} \binom{m}{k} k^n \tag{7}$$

(Temme 1996, p. 21), which has the property $\mathcal{S}_n^{(n)} = 1$.

Evaluating the sum in (5) may cause severe cancellation of leading digits. For example, the sum for $n = 100$ and $x = 5000$ is[1]

$$F_{n,1}(x) = 0.51721318855009723..., \tag{8}$$

but the terms in the sum (divided by $n!$) range in magnitude from 10^{-2327} (for $j = 1$) to 10^{27} (for $j = 86$), and the alternating signs should provide the proper answer. That is why we now derive an asymptotic expansion, of which only the first few terms will already give sharp approximations for $F_{n,\alpha}(x)$. To obtain the expansion, we first derive an integral representation for $F_{n,\alpha}(x)$ in terms of a contour integral in the complex plane. We start with the representation

$$F_{n,\alpha}(x) = 1 - G_{n,\alpha}(x), \tag{9}$$

where $G_{n,\alpha}(x)$ has the integral representation

$$G_{n,\alpha}(x) = \frac{1}{2\pi i} \int_{\mathcal{L}} \frac{e^{xs}}{s(s+1)(2^\alpha s+1)\cdots(n^\alpha s+1)}\, ds, \tag{10}$$

and $\mathcal{L}$ is a contour running in the positive (i.e., anti-clockwise) direction around the poles at $s = 0$ and at $s = -j^{-\alpha}, j = 1, 2, \ldots, n$. Taking into account the behavior of the integrand as $s \to \pm i\infty$, we can deform the finite contour into two vertical lines, one with $\Re s < -n^{-\alpha}$ and one with $\Re s > 0$. The left line can be shifted to the left as far as we like, and its contribution vanishes as we shift it to $-\infty$. So we are left with one vertical line $\mathcal{L}$ with $\Re s > 0$, and the integration runs from $-i\infty$ to $+i\infty$.

By shifting the vertical line $\mathcal{L}$ in (10) to the left, across the $n + 1$ poles, we obtain a sum of residues. The residue at $s = 0$ equals 1, and at $s = -j^{-\alpha}, j = 1, 2, \ldots, n$, it equals $-e^{-xj^{-\alpha}}/\Pi_{j,n,\alpha}$. In this way we obtain the sum in (4). On the other hand, when $x = 0$ we can shift the vertical line $\mathcal{L}$ in (10) to the right, as far as we like, and its contribution vanishes as we shift it to $+\infty$. This

[1] George Marsaglia mentioned on June 30, 1999 this case in Article 47714 of the Internet newsgroup sci.math.num-analysis , "A tough numerical problem", and wondered if it would be possible to evaluate, to at least 6-digit (single-precision) accuracy, the sum for the given values of x and n in (5) using IEEE standard double-precision arithmetic.

shows that $G_{n,\alpha}(0) = 0$, which confirms (6).

When $\alpha = 0$ we have

$$G_{n,0}(x) = \frac{1}{2\pi i}\int_{\mathcal{L}} \frac{e^{xs}}{s(s+1)^n}\,ds, \tag{11}$$

and this integral can be written in terms of the incomplete gamma function. We have (see (Temme 1996, pp. 282–283))

$$G_{n,0}(x) = P(n,x), \quad F_{n,0}(x) = Q(n,x), \tag{12}$$

where the incomplete gamma functions P and Q are defined by

$$P(a,z) = \frac{1}{\Gamma(a)}\int_0^z t^{a-1}e^{-t}\,dt,, \quad Q(a,z) = \frac{1}{\Gamma(a)}\int_z^\infty t^{a-1}e^{-t}\,dt. \tag{13}$$

When $\alpha = 1$ we can write the integrand of (10) in the form

$$\frac{e^{xs}}{s(s+1)(2s+1)\cdots(ns+1)} = \frac{e^{sx}}{s^{n+1}}\,\frac{\Gamma(1+1/s)}{\Gamma(n+1+1/s)}. \tag{14}$$

3 An asymptotic expansion of $F_{n,\alpha}(x)$

For obtaining an asymptotic representation we apply the saddle point method (see (Wong 2001, Ch. 2)). We write

$$G_{n,\alpha}(x) = \frac{1}{2\pi i}\int_{\mathcal{L}} e^{\phi(s)}\frac{ds}{s}, \tag{15}$$

where

$$\phi(s) = xs - \ln\left((s+1)(2^\alpha s+1)\cdots(n^\alpha s+1)\right). \tag{16}$$

The saddle point, defined by the $s-$value for which $\phi'(s) = 0$, is governed by the equation

$$x = \frac{1}{s+1} + \frac{2^\alpha}{2^\alpha s+1} + \ldots + \frac{n^\alpha}{n^\alpha s+1}. \tag{17}$$

It is not difficult to verify that, if $x > 0$, this equation has a unique solution s_0 in the interval $(-1/n^\alpha, \infty)$; see Figure 1. The curve represents the right-hand side of (17), and it cuts the vertical axis at a point indicated by

$$\mu_{n,\alpha} = \mathbb{E}(Z_{n,\alpha}) = 1 + 2^\alpha + \cdots + n^\alpha. \tag{18}$$

If $x > \mu_{n,\alpha}$ the saddle point is negative, as shown in Figure 1. If $x = \mu_{n,\alpha}$ the pole at the origin in (15) coincides with the saddle point. As explained in (Wong 2001, Ch. 7) we can use an error function to handle this case. We have the approximation

$$s_0 = -\frac{x-\mu_{n,\alpha}}{\mu_{n,2\alpha}}\left(1+o(1)\right), \quad x \to \mu_{n,\alpha}. \tag{19}$$

More details on this relation follow from §4.1, in particular by inverting the expansion in (35).

We transform the integral into a standard form for applying Laplace's method (Wong 2001, Ch. 2) by substituting

$$\phi(s) - \phi(s_0) = \tfrac{1}{2}w^2 \tag{20}$$

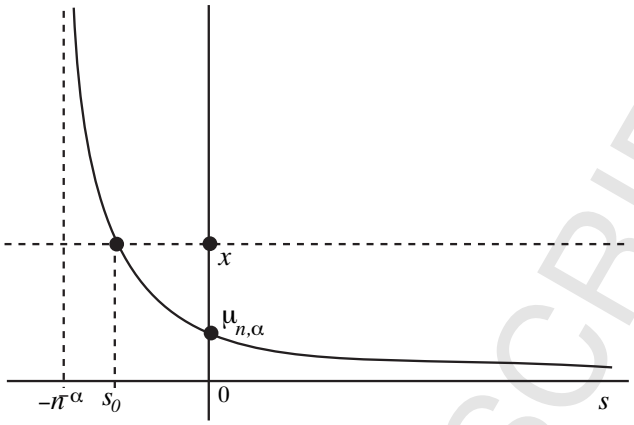


Figure 1: The saddle point s_0 is the solution of the equation (17). The curve represents the right-hand side of (17)

and prescribe that w has the sign of $s - s_0$ for real values of s with $s > -n^{-\alpha}$.

We obtain from (15)

$$G_{n,\alpha}(x) = \frac{e^{\phi(s_0)}}{2\pi i} \int_{-i\infty}^{i\infty} e^{\frac{1}{2}w^2} f(w) \frac{dw}{w-\xi}, \tag{21}$$

where

$$f(w) = \frac{w-\xi}{s} \frac{ds}{dw}, \tag{22}$$

and ξ is the $w-$value that corresponds with $s = 0$. That is, since $\phi(0) = 0$,

$$\tfrac{1}{2}\xi^2 = \phi(0) - \phi(s_0) = -\phi(s_0) \quad \Longrightarrow \quad \xi = \sqrt{-2\phi(s_0)}, \tag{23}$$

where ξ has the sign as prescribed in the mapping in (20). Thus, ξ should have the sign of $-s_0$, or in other words, ξ has the sign of $(x - \mu_{n,\alpha})$.

Because $\phi(0) = 0$, and $\phi(s) \to +\infty$ as $s \to -n^{-\alpha}$ and $s \to +\infty$, it follows that $\phi(s_0)$ is non-positive. Also, $f(\xi) = 1$, as easily follows from (22) by applying l'Hôpital's rule.

We see that, when $s_0 > 0$ and we take the vertical line through s_0, the pole at $s = 0$ is at the left side of $\mathcal{L}$. We obtain the same situation in (21): if $s_0 > 0$, that is, $\xi < 0$, then the pole at $w = \xi$ is at the left of the saddle point at $w = 0$. To handle the integral in (21) for small values of ξ (when $x \sim \mu_{n,\alpha}$) we regularize the integrand by splitting off the pole at $w = \xi$ writing

$$\frac{f(w)}{w-\xi} = g(w) + \frac{1}{w-\xi}, \quad g(w) = \frac{f(w)-1}{w-\xi}. \tag{24}$$

This gives

$$G_{n,\alpha}(x) = \frac{e^{\phi(s_0)}}{2\pi i} \int_{-i\infty}^{i\infty} e^{\frac{1}{2}w^2} \frac{dw}{w-\xi} + \frac{e^{\phi(s_0)}}{2\pi i} \int_{-i\infty}^{i\infty} e^{\frac{1}{2}w^2} g(w) \frac{dw}{w-\xi}. \tag{25}$$

The first integral can be written as (by substituting $w = it$)

$$\frac{e^{\phi(s_0)}}{2\pi i}\int_{-i\infty}^{i\infty} e^{\frac{1}{2}w^2}\frac{dw}{w-\xi} = \frac{e^{\phi(s_0)}}{2\pi i}\int_{-\infty}^{\infty} e^{-\frac{1}{2}t^2}\frac{dt}{t+i\xi}. \tag{26}$$

When $\xi < 0$ this integral can be written as a complementary error function and a $w-$function; see Temme (2010)[2]. In this way, by using also (23),

$$\frac{e^{\phi(s_0)}}{2\pi i}\int_{-\infty}^{\infty} e^{-\frac{1}{2}t^2}\frac{dt}{t+i\xi} = \tfrac{1}{2}e^{\phi(s_0)}w\left(-i\xi/\sqrt{2}\right) = \tfrac{1}{2}\mathrm{erfc}\left(-\xi/\sqrt{2}\right). \tag{27}$$

The right-hand side is an entire function of ξ, and we will use this relation for all real ξ, although the starting point for evaluating the integral in (27) was $\xi < 0$.

The function g is analytic in the same domain as where f is analytic. We expand

$$g(w) = \sum_{k=0}^{\infty} c_k w^k. \tag{28}$$

This gives the asymptotic expansion

$$G_{n,\alpha}(x) \sim \tfrac{1}{2}\mathrm{erfc}(-\xi/\sqrt{2}) + \frac{e^{\phi(s_0)}}{\sqrt{2\pi}}\sum_{k=0}^{\infty} C_k, \quad C_k = (-1)^k\, 2^k\, (\tfrac{1}{2})_k\, c_{2k}. \tag{29}$$

For the function $F_{n,\alpha}(x)$ we obtain, using (9) and $1 - \frac{1}{2}\mathrm{erfc}\, z = \frac{1}{2}\mathrm{erfc}(-z)$, our main result:

$$F_{n,\alpha}(x) \sim \tfrac{1}{2}\mathrm{erfc}(\xi/\sqrt{2}) - \frac{e^{\phi(s_0)}}{\sqrt{2\pi}}\sum_{k=0}^{\infty} C_k. \tag{30}$$

There is no large parameter showing in the asymptotic series, but the coefficients contain the asymptotic information. The expansion holds for large values of n, uniformly with respect to $x \geq \delta$, where δ is a fixed positive number. For further information on this point we refer to §4.2.

4 Evaluating the coefficients

For evaluating the coefficients C_k in (30) we need the coefficients b_k in the expansion

$$s = \sum_{k=0}^{\infty} b_k w^k, \quad b_0 = s_0, \tag{31}$$

which follow from the inversion of (20). Let

$$\phi(s) = \sum_{k=0}^{\infty} \frac{a_k}{k!}(s-s_0)^k, \quad a_k = \phi^{(k)}(s_0), \tag{32}$$

[2]http://dlmf.nist.gov/7.7.i, http://dlmf.nist.gov/7.7.E2

with $a_1 = \phi'(s_0) = 0$. Then the first coefficients b_k are given by

$$b_0 = s_0, \quad b_1 = \frac{1}{\sqrt{a_2}}, \quad b_2 = \frac{-a_3}{6a_2^2}, \quad b_3 = \frac{5a_3^2 - 3a_2a_4}{72a_2^{7/2}},$$
$$b_4 = \frac{45a_2a_3a_4 - 40a_3^3 - 9a_2^2a_5}{1080a_2^5}. \tag{33}$$

By using (22) and (23) we find the coefficients c_{2k} of $g(w)$ in (28). We have for the coefficients C_0 and C_1:

$$C_0 = \frac{s_0\sqrt{a_2} + \xi}{s_0\sqrt{a_2}\,\xi},$$
$$C_1 = -\frac{12a_3a_2^3s_0\xi^3 + 24a_2^4\xi^3 + 5s_0^2\xi^3a_2^2a_3^2 - 3\xi^3s_0^2a_2^3a_4 + 24s_0^3a_2^{11/2}}{24s_0^3\xi^3a_2^{11/2}}. \tag{34}$$

4.1 Expansions of the coefficients near the mean

For small values of s_0 (or ξ) we need expansions in terms of one of these small parameters. First we need the expansion of x in powers of s_0 (which is the solution of (17)). We can expand

$$x = \mu_{n,\alpha} - \mu_{n,2\alpha}s_0 + \mu_{n,3\alpha}s_0^2 - \ldots = \sum_{k=0}^{\infty}(-1)^k\mu_{n,(k+1)\alpha}s_0^k. \tag{35}$$

Furthermore, we have

$$\phi'(s) = x - \sum_{j=1}^{n}\frac{j^\alpha}{j^\alpha s + 1}, \quad \phi^{(k)}(s) = (-1)^k(k-1)!\sum_{j=1}^{n}\frac{j^{\alpha k}}{(j^\alpha s+1)^k}, \quad k \geq 2, \tag{36}$$

and

$$\phi(s) = (x - \mu_{n,\alpha})\,s + \tfrac{1}{2}\mu_{n,2\alpha}s^2 - \tfrac{1}{3}\mu_{n,3\alpha}s^3 + \ldots. \tag{37}$$

Using (35) we obtain

$$\phi(s_0) = -\tfrac{1}{2}\mu_{n,2\alpha}s_0^2 + \tfrac{2}{3}\mu_{n,3\alpha}s_0^3 - \tfrac{3}{4}\mu_{n,4\alpha}s_0^4 + \ldots. \tag{38}$$

We obtain, using (37),

$$a_k = \phi^{(k)}(s_0) = (-1)^k\sum_{m=0}^{\infty}\frac{(m+k-1)!}{m!}(-s_0)^m\mu_{n,(k+m)\alpha}, \tag{39}$$

and, using (23) and (38),

$$\xi^2 = 2\sum_{m=2}^{\infty}\frac{m-1}{m}\mu_{n,m\alpha}(-s_0)^m. \tag{40}$$

This gives (recall that ξ has the sign of $-s_0$, cf. the definition of ξ in (23))

$$\xi = -\sqrt{\mu_{n,2\alpha}}\, s_0 \left(1 - \frac{2\mu_{n,3\alpha}}{3\mu_{n,2\alpha}} s_0 + \frac{27\mu_{n,4\alpha}\mu_{n,2\alpha} - 8\mu_{n,3\alpha}^2}{36\mu_{n,2\alpha}^2} s_0^2 + \dots\right). \tag{41}$$

By using (39) and (41) we can obtain expansions for the quantities C_k of (30). For example,

$$C_0 = \frac{1}{3}\frac{\mu_{n,3\alpha}}{\mu_{n,2\alpha}^{3/2}} - \frac{1}{12}\frac{9\mu_{n,2\alpha}\mu_{n,4\alpha} - 10\mu_{n,3\alpha}^2}{\mu_{n,2\alpha}^{5/2}} s_0 + O\left(s_0^2\right), \tag{42}$$

and

$$C_1 = \frac{1}{540}\frac{675\mu_{n,2\alpha}\mu_{n,3\alpha}\mu_{n,4\alpha} - 350\mu_{n,3\alpha}^3 - 324\mu_{n,2\alpha}^2\mu_{n,5\alpha}}{\mu_{n,2\alpha}^{9/2}} + O\left(s_0\right). \tag{43}$$

4.2 On the order of growth of the coefficients at the mean

The coefficients C_k of (30) constitute an asymptotic scale for large n. This means $C_{k+1} = o\left(C_k\right)$ as $n \to \infty$. We verify this property at the mean, that is, when $x = \mu_{n,\alpha}$, or $s_0 = 0$. Exact values of C_0 and C_1 are given in (42) and (43) with $s_0 = 0$.

To estimate these coefficients for large n we observe that for $\alpha > -1$ we have $\mu_{n,\alpha} = O\left(n^{\alpha+1}\right)$. It follows that for $\alpha \geq 0$

$$C_0 = O(n^{-1/2}), \quad C_1 = O(n^{-3/2}). \tag{44}$$

Considering C_k with higher k we see the same pattern as in C_0 and C_1: the numerator of C_k contains a sum of products $\mu_{n,2\alpha}^p \mu_{n,3\alpha}^q \cdots$ and all these terms are of the same order $O\left(n^{3(2k+1)\alpha+2k+1}\right)$. The denominator is of order $O\left(n^{3(2k+1)\alpha+3k+3/2}\right)$, which gives

$$C_k = O(n^{-k-1/2}), \quad k = 0, 1, 2, \dots. \tag{45}$$

This has been derived at the mean for $\alpha \geq 0$. When $-\frac{1}{2} < \alpha < 0$, $\mu_{n,2\alpha}$ is still growing and the verification that C_k constitutes an asymptotic scale can be slightly modified, although other $\mu_{n,j\alpha}$ may have order $O(1)$.

5 Special cases

We shall now discuss some special cases related to specific values of α.

5.1 Central limit theorem ($\alpha \geq -\frac{1}{2}$)

We first consider those values of α for which $Z_{n,\alpha}$ falls within the domain of attraction of the CLT. Let $\xrightarrow{d}$ denote convergence in distribution, and let $N(0,1)$ be a standard normal random variable. Then, for $\alpha \geq -\frac{1}{2}$,

$$\frac{Z_{n,\alpha} - \mu_{n,\alpha}}{\sigma_{n,\alpha}} \xrightarrow{d} N(0,1), \tag{46}$$

as $n \to \infty$, where $\mu_{n,\alpha} = 1 + 2^\alpha + \cdots + n^\alpha$ and $\sigma_{n,\alpha}^2 = 1 + 2^{2\alpha} + \cdots + n^{2\alpha}$. This follows from the CLT with Lyapunov conditions (see (Chung 2001, p. 211)). Indeed, for $S_n = \sum_{j=1}^n X_n$, $s_n =$

$(\sum_{j=1}^{n} \mathrm{Var}(X_j))^{3/2}$, it is known that if

$$\sum_{j=1}^{n} \mathbb{E}(|X_j|^3)/s_n \longrightarrow 0, \tag{47}$$

then $S_n/s_n \xrightarrow{d} N(0,1)$. With $X_j = j(Y_j - 1)$, (47) is equivalent with

$$T_{n,\alpha} := \frac{\sum_{j=1}^{n} j^{3\alpha}}{(\sum_{j=1}^{n} j^{2\alpha})^{3/2}} \longrightarrow 0. \tag{48}$$

The latter is only true for $\alpha \geq -\frac{1}{2}$, which gives (46).

An alternative proof of (46) can be obtained from expanding the Laplace transform. That is,

$$\begin{aligned}\prod_{j=1}^{n} \frac{\exp(s\mu_{n,\alpha}/\sigma_{n,\alpha})}{1 + sj^{\alpha}/\sigma_{n,\alpha}} &= \exp(s\mu_{n,\alpha}/\sigma_{n,\alpha}) \exp\left(-\sum_{j=1}^{n}\left(s\frac{j^{\alpha}}{\sigma_{n,\alpha}} - s^2\frac{j^{2\alpha}}{2\sigma_{n,\alpha}^2} + s^3\frac{j^{3\alpha}}{3\sigma_{n,\alpha}^3} - \dots\right)\right) \\ &= \exp\left(\tfrac{1}{2}s^2 - T_{n,\alpha}s^3 + \dots\right).\end{aligned} \tag{49}$$

The CLT thus applies, and gives

$$F_{n,\alpha}(x) \approx 1 - \Phi\left(\frac{x - \mu_{n,\alpha}}{\sigma_{n,\alpha}}\right), \tag{50}$$

where $\Phi(x)$ is the normal distribution function

$$\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-t^2/2}\, dt = \tfrac{1}{2}\mathrm{erfc}(-x/\sqrt{2}). \tag{51}$$

The approximation (50) is in particular useful for large values of n, with x close to $\mu_{n,\alpha}$. Observe that from (19) and (41) we can derive

$$\xi \sim \frac{x - \mu_{n,\alpha}}{\sigma_{n,\alpha}}, \quad \text{as} \quad x \sim \mu_{n,\alpha}, \tag{52}$$

which shows that the term with the complementary error function in (30) corresponds with the approximation in (50), when $x \sim \mu_{n,\alpha}$.

We shall now compare (50) with our uniform expansion (30), for which we consider $\alpha = 1$, $n = 100$, $x = 5000$ and compare the approximations in (50) and (30) with the value given in (8). Then (see (18)) $\mu = \mu_{n,1} = 5050$ and $\sigma_{n,1} = 581.6786\ldots$. This gives for (50) the CLT approximation

$$F_{n,1}(x) \approx 0.53425. \tag{53}$$

Next we consider (30). We have $s_0 = 0.0001494392\ldots$ and $\xi = -0.08627917\ldots$, which gives (by taking only the error function)

$$F_{n,1}(x) \approx 0.53438. \tag{54}$$

Taking into account the first term of the expansion we have $C_0 = 0.0432412\ldots$, which gives

$$F_{n,1}(x) \approx 0.51719. \tag{55}$$

Comparing this with (8) we observe a relative error 0.42×10^{-4}. Including the contribution from $C_1 = -0.5616593 \times 10^{-4}$, we obtain

$$F_{n,1}(x) \approx 0.51721340, \tag{56}$$

with relative error 0.42×10^{-6}.

More numerical details will be given in Section 6.

5.2 Erlang distribution ($\alpha = 0$)

Remember that for $\alpha = 0$ the random variable $Z_{n,\alpha}$ is in fact an i.i.d. sum of exponentials, and the distribution function is in fact the Erlang distribution, which can be expressed in terms of the incomplete gamma function, see (12). For $\alpha = 0$, we have

$$s_0 = \frac{1}{\lambda} - 1, \quad \phi(s_0) = n(1 - \lambda + \ln \lambda), \quad \lambda = \frac{x}{n}, \tag{57}$$

and

$$\xi = \sqrt{2n(-1 + \lambda - \ln \lambda)}, \quad a_2 = n\lambda^2, \tag{58}$$

where ξ has the sign of $\lambda - 1$. This gives the same coefficients C_k (up to scaling) as derived in Temme (1979) for the incomplete gamma functions. The first term in the expansion (30) agrees with the first term in the Lugannani and Rice formula for the sums of i.i.d. random variables.

5.3 Gumbel distribution ($\alpha = -1$)

For $\alpha = -1$ we have that

$$G_{n,-1}(x) = 1 + \sum_{j=1}^{n} \binom{n}{j} (-1)^j e^{-xj} = (1 - e^{-x})^n, \tag{59}$$

and in fact, $Z_{n,-1}$ is equal in distribution to $\max\{Y_1, Y_2, \ldots, Y_n\}$, with $Y_1, Y_2, \ldots$ the i.i.d. unit mean exponentials, as before. From extreme value theory we know that the distribution of $Z_{n,-1}$ then converges to a *double exponential* or Gumbel distribution. That is, as $n \to \infty$,

$$Z_{n,-1} - \log n \xrightarrow{d} W, \tag{60}$$

with $\mathbb{P}(W < x) = \exp(-e^{-x})$, $x \in \mathbb{R}$. For the Laplace transform we find that

$$\lim_{n\to\infty} \nu_{n,-1}(s) e^{s \sum_{j=1}^{n} j^{-1}} = \lim_{n\to\infty} \prod_{j=1}^{n} \frac{e^{s/j}}{1 + s/j} = e^{\gamma s} \Gamma(s+1), \tag{61}$$

with $\gamma = \lim_{n\to\infty} (\sum_{j=1}^{n} j^{-1} - \log n)$.

Table 1: Values of $F_{n,\alpha}$ for $n = 10$ and $x = 0.99\mu_{n,\alpha}$, CLT is obtained from (50), erfc from (30) (without terms) and δ_0, δ_2 are the relative errors by taking the terms C_0 and C_0, C_1 in the expansion.

α	x	$F_{n,\alpha}$	CLT	erfc	δ_0	δ_2
-2.0	1.5342701	0.38907739	0.50594346	0.50597034	0.91e-2	0.60e-2
-1.5	1.9753831	0.40553885	0.50727377	0.50730961	0.13e-1	0.42e-2
-1.0	2.8996786	0.43228537	0.50938537	0.50943135	0.12e-1	0.53e-3
-0.5	4.9707880	0.46046846	0.51170255	0.51174833	0.35e-2	0.82e-3
0.0	9.9000000	0.47050158	0.51261356	0.51265584	0.37e-4	0.11e-4
0.5	22.243595	0.46580821	0.51208461	0.51212754	0.54e-3	0.22e-4
1.0	54.450000	0.45853324	0.51118113	0.51122320	0.13e-2	0.99e-4
1.5	141.24559	0.45179788	0.51034755	0.51038785	0.22e-2	0.21e-3
2.0	381.15000	0.44578999	0.50964906	0.50968747	0.31e-2	0.36e-3

5.4 Kolmogorov distribution ($\alpha = -2$)

For $\alpha = -2$ we obtain

$$\Pi_{j,\infty,-2} = \prod_{k=1,k\neq j}^{\infty} (1 - (j/k)^2) = -\frac{\cos \pi j}{2} \tag{62}$$

and hence

$$G_{\infty,-2}(x) = 1 - 2\sum_{j=1}^{\infty}(-1)^{j+1}e^{-xj^2} = \sum_{j=-\infty}^{\infty}(-1)^j e^{-xj^2}, \tag{63}$$

which is known as Kolmogorov's distribution. Using the Jacobi identity, (63) can be represented as an infinite product (see Chung (1982)):

$$G_{\infty,-2}(x) = \prod_{j=1}^{\infty} \frac{1 - e^{-jx}}{1 + e^{-jx}}. \tag{64}$$

The corresponding Laplace transform reads

$$\lim_{n\to\infty} \nu_{n,-2}(s) = \prod_{j=1}^{\infty} \frac{1}{1 + s/j^2} = \frac{\pi\sqrt{s}}{\sinh \pi\sqrt{s}}, \tag{65}$$

where the last equality follows from

$$\sinh z = z \prod_{j=1}^{\infty}(1 + \frac{z^2}{j^2\pi^2}). \tag{66}$$

See also Kac and Siegert (1947) and Biane et al. (2001) for more applications of the case $\alpha = -2$.

Table 2: Similar values as in Table 1, now for $n = 100$ and again $x = 0.99\mu_{n,\alpha}$.

α	x	$F_{n,\alpha}$	CLT	erfc	δ_0	δ_2
-2.0	1.6186341	0.38945105	0.50626942	0.50629932	0.91e-2	0.60e-2
-1.5	2.3887454	0.40771270	0.50877921	0.50883121	0.13e-1	0.41e-2
-1.0	5.1355037	0.44575400	0.51618014	0.51630710	0.12e-1	0.81e-3
-0.5	18.403708	0.50696945	0.53252552	0.53270653	0.19e-2	0.39e-3
0.0	99.000000	0.52669567	0.53982784	0.53996093	0.13e-5	0.31e-7
0.5	664.74825	0.52281707	0.53763922	0.53777378	0.17e-4	0.84e-7
1.0	4999.5000	0.51755849	0.53459179	0.53472196	0.43e-4	0.42e-6
1.5	40096.212	0.51287625	0.53196106	0.53208556	0.74e-4	0.10e-5
2.0	334966.50	0.50881793	0.52978242	0.52990111	0.11e-3	0.20e-5

Table 3: Similar values as in Table 1, now for $n = 100$, $x = 0.50\mu_{n,\alpha}$, $G_{n,\alpha}$ and the corresponding CLT and expansion (29).

α	x	$G_{n,\alpha}$	CLT	erfc	δ_0	δ_2
-2.0	0.8174920	0.1975542e-0	0.2159960e-0	0.1352230e-0	0.24e-1	0.66e-3
-1.5	1.2064370	0.6433725e-1	0.1355787e-0	0.4486680e-1	0.21e-1	0.43e-2
-1.0	2.5936888	0.4228630e-3	0.2125793e-1	0.3132675e-3	0.74e-2	0.85e-3
-0.5	9.2948020	0.1518863e-7	0.2242152e-4	0.1198045e-7	0.85e-3	0.35e-4
0.0	50.000000	0.3200065e-9	0.2866516e-6	0.2562014e-9	0.84e-5	0.31e-6
0.5	335.73144	0.1657144e-8	0.1154000e-5	0.1318627e-8	0.10e-3	0.17e-6
1.0	2525.0000	0.1908872e-7	0.7095496e-5	0.1508219e-7	0.21e-3	0.98e-6
1.5	20250.612	0.1545196e-6	0.3035652e-4	0.1214514e-6	0.31e-3	0.22e-5
2.0	169175.00	0.8285763e-6	0.9343075e-4	0.6486733e-6	0.41e-3	0.38e-5

6 Numerical experiments

In Table 1 we give numerical details of the asymptotic approximations. We take $n = 10$ for all cases and $x = 0.99\mu_{n,\alpha}$, these values being close to the mean. $F_{n,\alpha}$ is the value obtained from multi-precision arithmetic, CLT denotes the value obtained from (50) (also for values of $\alpha < -\frac{1}{2}$, where the CLT approach is no longer valid). The column with erfc gives the values obtained from (30) taking only the complementary error function, and δ_0, δ_2 are the relative errors taking also the terms C_0 and C_0, C_1 in the expansion.

In Table 2 we use $n = 100$ and again $x = 0.99\mu_{n,\alpha}$. In Table 3 we use $n = 100$ and $x = 0.50\mu_{n,\alpha}$, which is outside the mean values, but now for $G_{n,\alpha}$ and the corresponding CLT and asymptotic expansion (29) (because $F_{n,\alpha}$ is almost equal to unity in some cases).

We observe that in Table 1 and Table 2 the values of the columns CLT and erfc are quite close, because x and n are close to the mean. In all cases the results become sharper when we add the terms C_0 and C_1 in the asymptotic expansion (30). The computations are done in Maple with 30-digits arithmetic, except for the computation of $F_{n,\alpha}(x)$, where for $n = 100$ we took 300 digits.

Acknowledgements

We thank the referee for helpful suggestions. We thank George Marsaglia for bringing the case $\alpha = 1$ to our attention. NMT acknowledges financial support from *Gobierno of Navarra*, Res. 07/05/2008, *Ministerio de Educación y Ciencia* (project MTM2006–09050), and *Ministerio de Ciencia e Innovación* (project MTM2009-11686). JvL is supported by an ERC Starting Grant.

References

P. Biane, J. Pitman, and M. Yor. Probability laws related to the Jacobi theta and Riemann zeta functions, and Brownian excursions. *Bull. Amer. Math. Soc. (N.S.)*, 38(4):435–465 (electronic), 2001. ISSN 0273-0979.

K. L. Chung. A cluster of great formulas. *Acta Math. Acad. Sci. Hungar.*, 39(1-3):65–67, 1982. ISSN 0001-5954.

K. L. Chung. *A Course in Probability Theory*. Academic Press Inc., San Diego, CA, third edition, 2001. ISBN 0-12-174151-6.

L. Holst. On the lengths of the pieces of a stick broken at random. *J. Appl. Probab.*, 17(3):623–634, 1980. ISSN 0021-9002.

M. Kac and A. J. F. Siegert. On the theory of noise in radio receivers with square law detectors. *J. Appl. Phys.*, 18:383–397, 1947.

J. F. C. Kingman. Martingales in the OK Corral. *Bull. London Math. Soc.*, 31(5):601–606, 1999. ISSN 0024-6093.

J. F. C. Kingman and S. E. Volkov. Solution to the OK Corral model via decoupling of Friedman's urn. *J. Theoret. Probab.*, 16(1):267–276, 2003. ISSN 0894-9840.

R. Lugannani and S. Rice. Saddle point approximation for the distribution of the sum of independent random variables. *Adv. in Appl. Probab.*, 12(2):475–490, 1980. ISSN 0001-8678.

T. Matsunawa. The exact and approximate distributions of linear combinations of selected order statistics from a uniform distribution. *Ann. Inst. Statist. Math.*, 37(1):1–16, 1985. ISSN 0020-3157.

N. M. Temme. The asymptotic expansion of the incomplete gamma functions. *SIAM J. Math. Anal.*, 10(4): 757–766, 1979. ISSN 0036-1410.

N. M. Temme. *Special functions: An Introduction to the Classical Functions of Mathematical Physics*. A Wiley-Interscience Publication. John Wiley & Sons Inc., New York, 1996. ISBN 0-471-11313-1.

N. M. Temme. Chapter 7, Error Functions, Dawson's and Fresnel Integrals. In F.W.J. Olver, D.W. Lozier, R.F. Boisvert, and Ch.W. Clark, editors, *NIST Handbook of Mathematical Functions*. Cambridge University Press, Cambridge, UK, 2010. ISBN ISBN-13:9780521192255.

R. Wong. *Asymptotic Approximations of Integrals*, volume 34 of *Classics in Applied Mathematics*. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 2001. ISBN 0-89871-497-4. Corrected reprint of the 1989 original.