This is a preprint of: "On the norming constants for normal maxima", Armengol Gasull, Maria Jolis, Frederic Utzet, *J. Math. Anal. Appl.*, vol. 422, 376–396, 2015.
DOI: [10.1016/j.jmaa.2014.08.025]

On the norming constants for normal maxima

ARMENGOL GASULL, MARIA JOLIS AND FREDERIC UTZET

Departament de Matemàtiques, Edifici C, Universitat Autònoma de Barcelona, 08193 Bellaterra (Barcelona) Spain. E-mails: gasull@mat.uab.cat, mjolis@mat.uab.cat, utzet@mat.uab.cat

Abstract. In a remarkable paper, Peter Hall [*On the rate of convergence of normal extremes*, J. App. Prob, **16** (1979) 433–439] proved that the supremum norm distance between the distribution function of the normalized maximum of n independent standard normal random variables and the distribution function of the Gumbel law is bounded by $3/\log n$. In the present paper we prove that choosing a different set of norming constants that bound can be reduced to $1/\log n$. As a consequence, using the asymptotic expansion of a Lambert W type function, we propose new explicit constants for the maxima of normal random variables.

Keywords: Gaussian law, extreme value theory, Lambert W function.

AMS classification: 60G70, 60F05, 62G32

1 Introduction

Let $X_1, \ldots, X_n$ be i.i.d. standard normal random variables and denote by M_n its maximum,

$$M_n = \max\{X_1, \ldots, X_n\}.$$

The normal law is in the domain of attraction for maxima of the Gumbel law, that is, there are sequences of real numbers $\{a_n,\ n \geq 1\}$ and $\{b_n,\ n \geq 1\}$ (the *norming –or normalizing– constants*) with $a_n > 0$ such that

$$\lim_n \frac{1}{a_n}\left(M_n - b_n\right) = G, \text{ in distribution,} \tag{1}$$

where G is a Gumbel random variable, with distribution function

$$\Lambda(x) = \exp\{-e^{-x}\},\ x \in \mathbb{R}. \tag{2}$$

Denote by $\Phi(x)$ the distribution function of a standard normal law and by $\phi(x)$ its density. The convergence (1) is equivalent that for every $x \in \mathbb{R}$,

$$\lim_n \Phi^n(a_n x + b_n) = \Lambda(x). \tag{3}$$

In a remarkable paper Peter Hall [7] proved that taking b_n^* such that

$$\frac{1}{\sqrt{2\pi}} \frac{1}{b_n^*} e^{-(b_n^*)^2/2} = \frac{1}{n} \quad \text{and} \quad a_n^* = 1/b_n^*,$$

it holds that for $n \geq 2$,

$$\frac{C'}{\log n} < \sup_{x\in\mathbb{R}} |\Phi^n(a_n^* x + b_n^*) - \Lambda(x)| < \frac{C}{\log n}, \tag{4}$$

with $C = 3$, and that the rate of convergence cannot be improved by choosing a different sequence of norming constants. In this way, Hall gives a precise quantification of the remark by Fihser and Tippet in the the seminal paper [5]: *From the normal distribution the limiting distribution is approached with extreme slowness*. Notice that if $2 \leq n \leq 20$, then $3/\log n > 1$, so the upper bound in (4) gives no information. It should also be remarked that Hall [7] points out that his constant C in (4) can be decreased to 0.91 when $n \geq 10^6$.

In the present paper we prove that taking

$$b_n = \Phi^{-1}\big(1 - \frac{1}{n}\big) \quad \text{and} \quad a_n^\circ = \frac{b_n}{1 + b_n^2}, \tag{5}$$

we have the following theorem.

Theorem 1. *Given $n_0 \geq 5$, for all $n \geq n_0$ it holds that*

$$\sup_{x\in\mathbb{R}} |\Phi^n\big(a_n^\circ x + b_n\big) - \Lambda(x)| < \frac{C(n_0)}{\log n},$$

with

$$C(n_0) = \begin{cases} 1, & \textit{when} \quad n_0 \leq 15 \\ \Big(\dfrac{2}{3b_{n_0}^2} + \dfrac{1}{\sqrt{e n_0}}\Big)\log(n_0) < 1 & \textit{when} \quad n_0 \geq 16. \end{cases}$$

Moreover $\lim_{n_0\to\infty} C(n_0) = 1/3$.

The above result is quite sharp because our numerical analysis shows that when n_0 moves in the range $[10^{20}, 10^{60}]$, then $C(n_0)$ cannot be taken smaller than 0.12, see Table 2. In Proposition 4 we give some bounds for $\{b_n^2\}$ that in particular prove that when $n_0 \geq 16$,

$$C(n_0) \leq \widetilde{C}(n_0) = \frac{1}{3}\frac{1}{1 - \dfrac{\log(4\pi\log n_0))}{2\log n_0}} + \frac{\log n_0}{\sqrt{e n_0}},$$

obtaining explicit and simple computable upper bounds for $C(n_0)$. To have an idea of how $C(n_0)$ and $\widetilde{C}(n_0)$ change with n_0 we present some values in Table 1.

n_0	16	30	50	10^2	10^4	10^6	10^{10}	10^{20}	10^{100}
$C(n_0)$	0.90	0.75	0.67	0.60	0.45	0.41	0.38	0.36	0.34
$\widetilde{C}(n_0)$	1.10	0.82	0.72	0.63	0.45	0.41	0.38	0.36	0.34

Table 1. Several upper approximations for $C(n_0)$ and $\widetilde{C}(n_0)$.

From a practical point of view, in order to have explicit expressions of the constants, it is suggested the following asymptotic equivalents to the norming constants b_n^* and a_n^*, respectively (Hall [7, Diplay (4)]):

$$\beta_n^* = (2\log n)^{1/2} - \log(4\pi\log n)/\big(2(2\log n)^{1/2}\big) \quad \text{and} \quad \alpha_n^* = 1/\beta_n^*.$$

(It is also proposed $\alpha_n^* = (2\log n)^{-1/2}$, see, for example, Resnick [9, pp. 71–72]). The expression of β_n^* is easily deduced by observing that b_n^* can be expressed in terms of the Lambert W function (Corless *et al.* [3]) and its well known asymptotics, see Section 5.

However, in view of Theorem 1 and our numerical computations (see again Table 2), on the one hand, it seems sensible to approach accurately b_n, rather than b_n^*, and we propose the constant

$$\beta_n = \left(\log\big(n^2/(2\pi)\big) - \log\log\big(n^2/(2\pi)\big) + \frac{\log\big(\log(n^2)+1/2\big) - 2}{\log\big(n^2/(2\pi)\big)}\right)^{1/2}, \tag{6}$$

that, as we will see, satisfies

$$b_n = \beta_n + O\left(\frac{(\log\log n)^2}{(\log n)^{5/2}}\right), \quad n \to \infty.$$

On the other hand, by Remark 10 and once more Table 2, it seems convenient to use

$$\alpha_n = \frac{\beta_n}{1+\beta_n^2},$$

rather than $1/\beta_n$. The expression of β_n is derived from an asymptotic expansion of b_n using an approximation to Mills ratio by rational functions (see Subsection 3.2), an extension of the asymptotics of the Lambert function to a more general class of functions (Subsection 5.1) and a final refinement motivated by some numerical computations (Subsection 5.3).

The paper is organized in the following way: In Section 2 there are recalled some elementary facts about Extreme Value Theory for normal random variables, and there are presented graphical and numerical comparative studies of the performance of the constants a_n^* and b_n^* versus a_n° and b_n, and other proposals. In Section 3 there are presented some technical preliminary results needed in next sections. Section 4 is devoted to proof of Theorem 1. Finally, in Section 5, new explicit expressions of the norming constant are given.

2 Extreme value theory for the normal law

By classical Extreme Value Theory, the norming constant b_n in (1) can be taken, and we take, in agreement with the notation (5),

$$b_n = \Phi^{-1}(1 - n^{-1}), \tag{7}$$

The constant a_n can be chosen to be

$$a_n = A(b_n), \tag{8}$$

where A is an auxiliary function corresponding to Φ (see, for example, Resnick [9, Proposition 1.11]). Auxiliary functions are not unique though they are asymptotically equal. Moreover, under certain conditions, an auxiliary function is (see again Resnick [9, Proposition 1.11]) the quotient of the survival function (one minus the distribution function) and the density function, that is,

$$A_{\mathcal{C}}(x) = \frac{1 - \Phi(x)}{\phi(x)}, \tag{9}$$

which is called the Mills ratio. Since this function is expressed in terms of the distribution function and the density function, and does not depend on any other computation, we call it the *canonical* auxiliary function. We should remark that from the standard proof of the convergence (1) it is not deduced that the constants b_n and $A_{\mathcal{C}}(b_n)$ produce more accurate results than other constants computed with other auxiliary functions or other ways.

To find manageable expression of the constants it is used a property of the convergence in law adapted to this context:

Property 2. *With the preceding notations, if the sequences* $\{a'_n,\, n \geq 1\}$ *and* $\{b'_n,\, n \geq 1\}$ *satisfy*

$$\lim_n \frac{a_n}{a'_n} = 1 \quad \textit{and} \quad \lim_n \frac{b_n - b'_n}{a_n} = 0,$$

then

$$\lim_n \frac{1}{a'_n}\big(M_n - b'_n\big) = G \textit{ in distribution.}$$

Moreover, it is very useful the following property that involves the use of the norming constants of a simpler distribution function right tail equivalent to Φ:

Property 3. *Let F be a distribution function right tail equivalent to Φ, that means,*

$$\lim_{x\to\infty} \frac{1 - \Phi(x)}{1 - F(x)} = 1.$$

Then the norming constants of F and Φ can be taken equal.

Thanks to the well known asymptotics of the Mills ratio,

$$\lim_{x\to\infty} \frac{1 - \Phi(x)}{\dfrac{1}{\sqrt{2\pi}}\,\dfrac{1}{x}\, e^{-x^2/2}} = 1,$$

we can consider a distribution function F such that there is some x_0, such that for $x > x_0$,

$$F(x) = 1 - \frac{1}{\sqrt{2\pi}} \frac{1}{x} e^{-x^2/2}, \tag{10}$$

and we deduce other possible constants: b_n^* is given by

$$b_n^* = F^{-1}(1 - n^{-1}),$$

or, equivalently, b_n^* verifies

$$\frac{1}{\sqrt{2\pi}} \frac{1}{b_n^*} e^{-(b_n^*)^2/2} = \frac{1}{n}. \tag{11}$$

On the other hand, the canonical auxiliary function associated to F is

$$A_{\mathcal{F}}(x) = \frac{x}{1 + x^2}. \tag{12}$$

We call this auxiliary function $A_{\mathcal{F}}$ because this was the election of Fisher and Tippett [5]. Note that $A_{\mathcal{C}}$ and $A_{\mathcal{F}}$ are asymptotically equivalent. In our early notations (5), we take

$$a_n^\circ = A_{\mathcal{F}}(b_n).$$

Furthermore, it is typical to use a simpler function asymptotically equivalent to both $A_{\mathcal{F}}$ and $A_{\mathcal{C}}$ given by

$$A_{\mathcal{H}}(x) = \frac{1}{x}.$$

We write

$$a_n^* := A_{\mathcal{H}}(b_n^*),$$

and we call a_n^* and b_n^* the *Hall's constants*. Hall didn't introduce such constants, that are classical (indeed, b_n^* was proposed by Fisher and Tippett [5]), but as we commented, Hall [7] proved the rate of convergence (4) with these constants. However, numerical studies show that other norming constants give more accurate results that Hall's ones. In Figure 1 there is a plot of the Gumbel density and the density of the random variables

$$Y_n^* = \frac{1}{a_n^*}(M_n - b_n^*) \quad \text{and} \quad Y_n = \frac{1}{a_n^\circ}(M_n - b_n).$$

for $n = 100$.

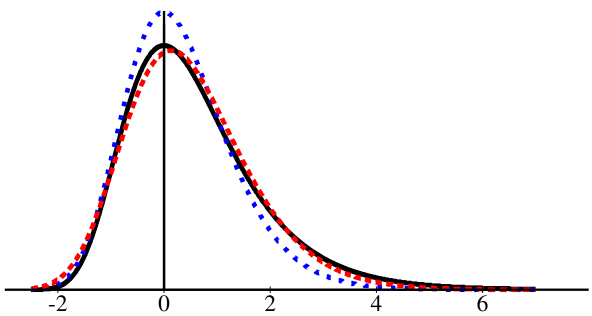

Figure 1. Gumbel density and density of the màximum of 100 standard Gaussian random variables with different norming constants. Solid line: Gumbel density. Dotted blue line: Density of Y_n^*. Dashed red line: Density of Y_n.

In order to assess the velocity of convergence, we numerically compute an approximation of the distances

$$\sup_{x \in \mathbb{R}} |\Phi^n(A(b_n)\, x + b_n) - \Lambda(x)| \log(n),$$

Used in	Constants		n 10	10^3	10^{10}	10^{30}	10^{50}	10^{60}
Theorem 1	b_n	$a_n^\circ = A_{\mathcal{F}}(b_n)$	0.0420	0.1049	0.1191	0.1208	0.1208	0.1207
Proposition 9	b_n	$A_{\mathcal{H}}(b_n)$	0.3117	0.2752	0.2552	0.2465	0.2443	0.2437
	b_n	$A_{\mathcal{C}}(b_n)$	0.1201	0.1260	0.1245	0.1224	0.1217	0.1215
Fisher & Tippett	b_n^*	$A_{\mathcal{F}}(b_n^*)$	0.2331	0.2268	0.1997	0.1945	0.1938	0.1936
Hall	b_n^*	$a_n^* = A_{\mathcal{H}}(b_n^*)$	0.3546	0.3650	0.3461	0.3354	0.3324	0.3316

Table 2. Approximate distance $\sup_{x\in\mathbb{R}} |\Phi^n(A(b_n)\,x + b_n) - \Lambda(x)|\log(n)$ for different sample size n and different sets of norming constants.

for several values of n, and for different auxiliary functions A, and also with Hall's constants a_n^* and b_n^*, and with the constants proposed by Fisher and Tippet [5] that are b_n^* and $A_{\mathcal{F}}(b_n^*)$. Those approximations are obtained computing numerically the maxima of the corresponding functions. We have used Maple. The results are given in Table 2.

Our purpose is to get theoretical explanations of those numerical results and our main result is Theorem 1 given in the Introduction. We restrict our study to the cases where $a_n = a_n^\circ = A_{\mathcal{F}}(b_n)$ and $a_n = A_{\mathcal{H}}(b_n)$, and we omit the case $a_n = A_{\mathcal{C}}(b_n)$: the reasons for that omission are the following: First, numerically and analytically $A_{\mathcal{F}}(b_n)$ is much simpler that the $A_{\mathcal{C}}(b_n)$; second, Table 2 suggests that the performances of both $a_n = A_{\mathcal{F}}(b_n)$ and $a_n = A_{\mathcal{C}}(b_n)$ are very similar; and finally, the study of $A_{\mathcal{C}}(b_n)$ has its own details and tricks, and its study would enlarge significantly the paper.

3 Preliminary results

This section is divided in two parts. In the first one we prove a couple of properties of the sequence $\{b_n\}$ which are needed in the sequel. In the second part we introduce the reciprocal of the canonical auxiliary function that allows to express in a convenient way the difference $\Phi^n(a_n x + b_n) - \Lambda(x)$.

3.1 Bounds for b_n

We prove that the bounds for $(b_n^*)^2$ given by Hall [7, display (2)]:

$$2\log n - \log(4\pi \log n) < (b_n^*)^2 < 2\log n, \tag{13}$$

are also satisfied for b_n^2.

Proposition 4. *For each $n \geq 2$ the following inequalities hold:*

$$2\log n - \log(4\pi \log n) < b_n^2 < 2\log n. \tag{14}$$

Proof. First of all, observe that for $n = 2$ we have that $b_2 = 0$, while $2\log 2 - \log(4\pi\log 2) < 0$ and $2\log 2 > 0$. So, we consider the case $n \geq 3$. To see the right hand side inequality in (14), we will prove that for $n \geq 3$,

$$1 - \frac{1}{n} < \Phi\big(\sqrt{2\log n}\,\big).$$

By the change of variables $y = \sqrt{2\log n}$, this inequality is equivalent to

$$1 - e^{-y^2/2} < \Phi(y),$$

for $y \geq \sqrt{2\log 3} \approx 1.14823$. This is the same that

$$\int_y^\infty \frac{1}{\sqrt{2\pi}}\, e^{-x^2/2}dx < \int_y^\infty x\, e^{-x^2/2}dx,$$

for $y \geq \sqrt{2\log 3}$. And this inequality is clear because $\frac{1}{\sqrt{2\pi}} \approx 0.3989$.

In order to prove the inequality on the left-hand side of (14), the argument is similar. First, the function $H(z) := 2\log z - \log\big(4\pi \log z)\big),\ z > 1$, is strictly increasing, being negative for $z = 3$ and $z = 4$, and positive for $z = 5$. So we will prove the inequality for $n \geq 5$. Define

$$h(z) := \sqrt{2\log z - \log(4\pi \log z)} = \sqrt{H(z)}.$$

It is clear that h is strictly increasing and maps each interval $[n, \infty)$ into $[h(n), \infty)$, for all $n \geq 5$. We must show that $\Phi(h(n)) < 1 - 1/n$ or, equivalently, that

$$\int_{h(n)}^{\infty} \frac{1}{\sqrt{2\pi}}\, e^{-x^2/2}\, dx > \frac{1}{n} = \int_n^{\infty} \frac{1}{y^2}\, dy.$$

By the change of variables $y = h^{-1}(x)$, the left hand side of the above inequality is equal to

$$\begin{aligned}
&\int_n^{\infty} \frac{1}{\sqrt{2\pi}}\, \exp\{-(2\log y - \log(4\pi \log y)/2\}\, h'(y)\, dy \\
&\quad = \int_n^{\infty} \frac{1}{\sqrt{2\pi}}\, \frac{1}{y}\, \sqrt{4\pi \log y}\, \frac{1}{2\sqrt{2\log y - \log(4\pi \log y)}} \Big[\frac{2}{y} - \frac{1}{y \log y}\Big]\, dy \\
&\quad = \int_n^{\infty} \frac{1}{\sqrt{2}}\, \frac{1}{y^2}\, \frac{2\log y - 1}{\sqrt{\big(2\log y - \log(4\pi \log y)\big)\log y}}\, dy.
\end{aligned}$$

To prove that this last term is greater than $\int_n^{\infty} \frac{1}{y^2}\, dy$ we should prove that for any $y \geq 5$,

$$\frac{2\log y - 1}{\sqrt{\big(2\log y - \log(4\pi \log y)\big)\log y}} > \sqrt{2}.$$

By the change of variables $u = \log y$ and after squaring the two terms of the inequality and some simplifications, we have to show that, for $u \geq \log 5$,

$$2u \log(4\pi u) > 4u - 1$$

that is the same that

$$g(u) := \log(4\pi u) - 2 + \frac{1}{2u} > 0,$$

for $u \geq \log 5$. And this is due to the fact that $g(\log 5) > 0$, and $g'(u) > 0$ for $u > \log 5$.

$\square$

In several parts of this work we will get the rate of convergence of $\Phi^n(a_n x + b_n)$ to $\Lambda(x)$ in terms of b_n^2, and later we translate it in terms of $\log n$. To this end, we use the following result:

Proposition 5. *For any $n_0 \geq 3$ and any $n > n_0$ the following inequality is satisfied:*

$$b_n^2 > K(n_0)\, \log n, \quad \textit{with} \quad K(n_0) = \frac{b_{n_0}^2}{\log n_0}.$$

Proof. Observe that the proposition is equivalent to say that the sequence $\{b_n^2/\log n\,, n \geq 3\}$ is increasing. Nevertheless, we will prove the result in an indirect way. Specifically, we will prove the following assertion: *For any $K \in (0, 2)$, the equation*

$$\Phi^{-1}(1 - \frac{1}{x}) - \sqrt{K \log x} = 0 \tag{15}$$

has a unique solution $x_0 > 1$, and $\Phi^{-1}(1 - \frac{1}{x}) - \sqrt{K \log x} > 0$ for any $x > x_0$.

Observe that if we take $K = K(n_0) = b_{n_0}^2/\log n_0$ (due to Proposition 4, the inequality $0 < K < 2$ is satisfied), the solution of the equation (15) is precisely $x_0 = n_0$, and so, the proposition will follow from the assertion.

Composing with Φ and doing the change of variables $y = \sqrt{K \log x}$, in order to prove the assertion, we must see that the equation

$$1 - e^{-\frac{y^2}{K}} = \Phi(y),$$

has a unique solution $y_0 > 0$ and the function $1 - e^{-\frac{y^2}{K}} - \Phi(y)$ is positive for $y > y_0$.

Define

$$M(y) := e^{-y^2/K} = \frac{2}{K} \int_y^\infty x e^{-\frac{x^2}{K}} \, dx$$

and

$$N(y) := 1 - \Phi(y) = \frac{1}{\sqrt{2\pi}} \int_y^\infty e^{-\frac{x^2}{2}} \, dx.$$

We have to prove that the equation $M(y) - N(y) = 0$ has a unique solution $y_0 > 0$ and that for $y > y_0$, $M(y) - N(y) < 0$. To prove this, we study the function $M(y) - N(y)$ for $y \geq 0$. Notice that

$$\operatorname{sign}\left(\frac{M'(y)}{N'(y)} - 1\right) = -\operatorname{sign}(M'(y) - N'(y)). \tag{16}$$

Therefore we introduce the function

$$g(y) := \frac{M'(y)}{N'(y)} = \frac{2\sqrt{2\pi}}{K} \, y \, e^{\frac{K-2}{2K} y^2},$$

and study the equation $g(y) = 1$, for $y \geq 0$. It is not difficult to show that it has exactly two solutions, y_1 and y_2, and that

$$0 < y_1 < \widetilde{y} := \sqrt{\frac{2-K}{K}} < y_2,$$

where $\widetilde{y}$ is the unique positive solution of $g'(y) = 0$.

Moreover $g(y) - 1$ is positive in (y_1, y_2) and negative on $[0, y_1) \cup (y_2, \infty)$. Using (16), we get that $M - N$ is increasing in $[0, y_1) \cup (y_2, \infty)$ and decreasing in (y_1, y_2).

Notice also that $M(0) - N(0) = 1 - 1/2 > 0$. Joining all the information we get that $M - N$ has at most one zero, $y = y^*$, in $(0, y_2]$ and, if exits, it is in $(y_1, y_2]$. In fact, since

$$M'(x) < N'(x) \quad \text{for} \quad x > y_2,$$

integrating both sides from y to infinity, we obtain that $M(y) - N(y) < 0$ for all $y > y_2$.

In short, $M - N$ has exactly one zero $y_0 = y^*$ in $[0, \infty)$, this zero belongs to the interval $(y_1, y_2]$ and moreover $M - N$ is negative for $y > y_0$. This fact finishes the proof of the assertion. $\square$

3.2 The canonical auxiliary function and its reciprocal

The canonical auxiliary function

$$A_{\mathcal{C}}(t) = \frac{1 - \Phi(t)}{\phi(t)}, \quad t > 0,$$

is known as Mills ratio and enjoys nice properties. In Baricz [1] or Gasull and Utzet [6] it is proved that it is completely monotone, that means, the derivatives alternate their signs: $A_{\mathcal{C}}(t) > 0$, and for $n \geq 1$,

$$(-1)^n A_{\mathcal{C}}^{(n)}(t) > 0, \text{ for } t > 0.$$

In particular $A_{\mathcal{C}}$ is strictly decreasing and strictly convex. It is also known how to construct two sequences of rational functions $\{P_n(t)/Q_n(t),\ n \geq 0\}$ with nonnegative integer coefficients and numerators and denominators with increasing degrees, such that for all $t > 0$,

$$\frac{Q_{n+1}(t)}{P_{n+1}(t)} < A_{\mathcal{C}}(t) < \frac{Q_n(t)}{P_n(t)},$$

see again [6] or [8]. We will use

$$\frac{t}{t^2+1} < A_{\mathcal{C}}(t) < \frac{t^2+2}{t^3+3t} < \frac{1}{t}. \tag{17}$$

Denote by $V(t)$ the reciprocal of the canonical auxiliary function

$$V(t) = \frac{1}{A_{\mathcal{C}}(t)} \tag{18}$$

Since $A_{\mathcal{C}}$ is strictly decreasing, V is strictly increasing. Moreover, the bounds for $A_{\mathcal{C}}$ give bounds for V. In particular, from (17), for $t > 0$,

$$t < V(t) < t + \frac{1}{t}. \tag{19}$$

The function $V(t)$ also provides a very useful way to express the function $\Phi^n(A(b_n)\,x + b_n) - \Lambda(x)$, which is a main object in this paper.

Proposition 6. *Set $b_n = \Phi^{-1}(1 - 1/n)$ and let $\{a'_n,\, n \geq 1\}$ be an arbitrary sequence of strictly positive numbers. For every $x \in \mathbb{R}$ it holds that*

$$\Phi^n(a'_n x + b_n) - \Lambda(x) = e^{-nS_n(x)}\Big(\Lambda\big(I_n(x)\big) - \Lambda(x)\Big) + \Lambda(x)\left(e^{-nS_n(x)} - 1\right), \tag{20}$$

where

$$I_n(x) = \int_{b_n}^{a'_n x + b_n} V(t)\,dt, \quad 0 < S_n(x) < \frac{C_n^2(x)}{2(1 - C_n(x))} \quad \text{and} \quad C_n(x) = \frac{1}{n}e^{-I_n(x)}. \tag{21}$$

Proof. Notice that for $y \in \mathbb{R}$,

$$1 - \Phi(y) = \exp\big(\log(1 - \Phi(y))\big) = \exp\Big\{\int_{-\infty}^{y} \frac{-\phi(t)}{1 - \Phi(t)}\,dt\Big\} = \exp\Big\{-\int_{-\infty}^{y} V(t)\,dt\Big\}.$$

Then

$$\begin{aligned} 1 - \Phi(a'_n x + b_n) &= \exp\Big\{-\int_{-\infty}^{b_n} V(t)\,dt\Big\}\exp\Big\{-\int_{b_n}^{a'_n x + b_n} V(t)\,dt\Big\} \\ &= \exp\Big\{\log\big(1 - \Phi(t)\big)\Big|_{-\infty}^{b_n}\Big\}\exp\Big\{-\int_{b_n}^{a'_n x + b_n} V(t)\,dt\Big\} = \frac{1}{n}\,e^{-I_n(x)}, \end{aligned} \tag{22}$$

where the last equality follows from the definition of b_n. Notice also that, by (22), $0 < \exp\big(-I_n(x)\big)/n < 1$.

The following formula is well-known and was already used by Hall in [7]. For $u \in (-1, 1)$,

$$\log(1 - u) = -u - r(u) \quad \text{with} \quad 0 \leq r(u) \leq \frac{u^2}{2(1-u)}.$$

Then, from (22),

$$\log\,\Phi^n(a'_n x + b_n) = n\,\log\Big(1 - \frac{1}{n}\,e^{-I_n(x)}\Big) = -e^{-I_n(x)} - nS_n(x), \tag{23}$$

where $S_n(x)$ satisfies the conditions given in (21). Hence,

$$\Phi^n(a'_n x + b_n) - \Lambda(x) = e^{-nS_n(x)}\Lambda\big(I_n(x)\big) - \Lambda(x),$$

and (20) follows adding and subtracting the term $e^{-nS_n(x)}\Lambda(x)$. $\square$

4 Proof of Theorem 1

To short the notations in this proof, we write a_n instead of $a_n^\circ = A_{\mathcal{F}}(b_n)$. We will consider separately the cases $x \geq 0$ and $x < 0$, where the rate of convergence is analyzed; later both cases are joined to get a global rate of convergence. In Subsection 4.4 there are some comments about the other norming constants $A_{\mathcal{H}}(b_n)$ and $A_{\mathcal{C}}(b_n)$.

4.1 Case $x \geq 0$.

We prove the following proposition:

Proposition 7. *Given $n_0 \geq 3$, for all $n \geq n_0$ it holds that*

$$\sup_{x \geq 0} |\Phi^n\big(a_n\, x + b_n\big) - \Lambda(x)| < \frac{C^+(n_0)}{\log n},$$

where

$$C^+(n_0) = \Big(\frac{1}{e\, b_{n_0}^2} + \frac{1}{2(n_0-1)}\Big) \log(n_0).$$

Proof. By Proposition 6, for $x \geq 0$, we have that

$$\begin{aligned} \big|\Phi^n(a_n x + b_n) - \Lambda(x)\big| &\leq e^{-nS_n(x)} \Big|\Lambda\big(I_n(x)\big) - \Lambda(x)\Big| + \Lambda(x)\big|e^{-nS_n(x)} - 1\big| \\ &\leq \Big|\Lambda\big(I_n(x)\big) - \Lambda(x)\Big| + \big|e^{-nS_n(x)} - 1\big|. \end{aligned} \tag{24}$$

We first study the term $\big|e^{-nS_n(x)} - 1\big|$. From (21), $0 < C_n(x) \leq 1/n$ (now $I_n(x) \geq 0$) and hence,

$$0 < S_n(x) < \frac{1}{2n(n-1)}.$$

Thus, since when $y \geq 0$, $1 - e^{-y} \leq y$, we get

$$\begin{aligned} \big|e^{-nS_n(x)} - 1\big| &= 1 - e^{-nS_n(x)} < nS_n(x) < \frac{1}{2(n-1)} \\ &= \frac{\log n}{2(n-1)} \frac{1}{\log n} \leq \frac{\log n_0}{2(n_0-1)} \frac{1}{\log n}, \end{aligned} \tag{25}$$

because the function $\log y/(y-1)$ is decreasing. Notice that the above inequality gives the second term of the right hand side of the statement.

To bound the other term we will study separately the cases whether $I_n(x) < x$ or $I_n(x) \geq x$. It can be seen that both situations occur.

1. Case $I_n(x) < x$. Here,

$$\begin{aligned} \Big|\Lambda\big(I_n(x)\big) - \Lambda(x)\Big| &= \Lambda(x) - \Lambda\big(I_n(x)\big) \leq \Lambda'\big(I_n(x)\big)\big(x - I_n(x)\big) \\ &= \Lambda\big(I_n(x)\big) e^{-I_n(x)}\big(x - I_n(x)\big) \leq e^{x - I_n(x)} e^{-x}\big(x - I_n(x)\big). \end{aligned} \tag{26}$$

where we have used that for $x > 0$, $\Lambda(x)$ is increasing, $\Lambda'(x)$ is decreasing, the Mean Value Theorem and that $\Lambda\big(I_n(x)\big) \leq 1$.

At this point observe that since $a_n = b_n/(b_n^2 + 1)$,

$$0 < x - I_n(x) \leq x - \int_{b_n}^{a_n x + b_n} t\, dt = x - \frac{(a_n x)^2}{2} - a_n b_n x \leq (1 - a_n b_n)x = \frac{x}{b_n^2 + 1},$$

where we utilize the bound $V(t) > t$ given in (19). Hence, plugging the above inequality in (26),

$$\Big|\Lambda\left(I_n(x)\right) - \Lambda(x)\Big| \le e^{\frac{x}{b_n^2+1}} e^{-x} \frac{x}{b_n^2+1} = e^{-\frac{b_n^2 x}{b_n^2+1}} \frac{x}{b_n^2+1}$$
$$= e^{-\frac{b_n^2 x}{b_n^2+1}} \frac{b_n^2 x}{b_n^2+1} \frac{1}{b_n^2} \le \max_{y\in[0,\infty)} \left\{e^{-y}y\right\} \frac{1}{b_n^2} = \frac{1}{eb_n^2}. \tag{27}$$

2. Case $I_n(x) \ge x$. Here,

$$\Big|\Lambda\big(I_n(x)\big) - \Lambda(x)\Big| = \Lambda\big(I_n(x)\big) - \Lambda(x) \le \Lambda'(x)\big(I_n(x) - x\big)$$
$$= \Lambda(x)e^{-x}\big(I_n(x) - x\big) \le e^{-x}\big(I_n(x) - x\big), \tag{28}$$

where we have used the same properties as above. We proceed also as in the previous situation, and moreover we use that when $y \ge 0$, then $\log(1+y) \le y$. Given that $V(t) < t + 1/t$ (see (19)),

$$0 \le I_n(x) - x \le \frac{1}{2}\, a_n^2 x^2 + a_n b_n x + \log\left(\frac{a_n}{b_n}\, x + 1\right) - x$$
$$\le \frac{b_n^2}{2(b_n^2+1)^2}\, x^2 + \frac{b_n^2}{b_n^2+1}\, x + \frac{1}{b_n^2+1}\, x - x = \frac{b_n^2}{2(b_n^2+1)^2}\, x^2 < \frac{1}{2b_n^2}\, x^2. \tag{29}$$

Hence, from (28),

$$\Big|\Lambda\big(I_n(x)\big) - \Lambda(x)\Big| \le \frac{x^2}{2}\, e^{-x} \frac{1}{b_n^2} < \frac{2}{e^2 b_n^2}. \tag{30}$$

The proposition follows joining (25), (27) and (30), using that $1/e > 2/e^2$ and applying Proposition 5. $\square$

4.2 Case $x < 0$.

Proposition 8. *Given $n_0 \ge 3$, for all $n \ge n_0$ it holds that*

$$\sup_{x<0} \left|\Phi^n\big(a_n\, x + b_n\big) - \Lambda(x)\right| < \frac{C^-(n_0)}{\log n},$$

where

$$C^-(n_0) = \begin{cases} 1, & \textit{when} \quad n_0 \le 15 \\ \Big(\dfrac{2}{3b_{n_0}^2} + \dfrac{1}{\sqrt{e}n_0}\Big)\log(n_0) & \textit{when} \quad n_0 \ge 16. \end{cases}$$

Proof. First notice that for $x < 0$, $\Lambda(x) < \Lambda(0) < 1/e$, and

$$\Phi^n\big(a_n x + b_n\big) < \Phi^n\big(b_n\big) = \Big(1 - \frac{1}{n}\Big)^n < \frac{1}{e}.$$

Hence, for $3 \le n \le 15$,

$$\sup_{x<0}\Big|\Lambda(x) - \Phi^n\big(a_n x + b_n\big)\Big| < \frac{1}{e} < \frac{1}{\log n}. \tag{31}$$

From now on we assume $n \ge 16$. For convenience, we divide the values of x according whether

$$x \in (-\infty, -b_n/a_n), \quad x \in [-b_n/a_n, -1.25\log b_n] \quad \text{or} \quad x \in (-1.25\log b_n, 0).$$

Notice that for $n \ge 16$, $-b_n/a_n < -1.25\log b_n$. We remark that in our approach the choice of the point $-1.25\log b_n$ is essential to obtain sharp bounds.

1. Case $x \in (-\infty, -b_n/a_n)$. Here we will bound separately $\Lambda(x)$ and $\Phi^n(a_n x + b_n)$. Since $a_n x + b_n < 0$, on the one hand, $\Phi\big(a_n x + b_n\big) < \Phi(0) = 0.5$, and therefore

$$0 < \Phi^n(a_n x + b_n) < \frac{1}{2^n}.$$

On the other hand,

$$0 < \Lambda(x) < \Lambda\left(-\frac{a_n}{b_n}\right) = \Lambda\left(-(b_n^2+1)\right) = \exp\left\{-e^{b_n^2+1}\right\}.$$

Joining the above inequalities we obtain that for $x \in (-\infty, -b_n/a_n)$,

$$\Big|\Lambda(x) - \Phi^n\big(a_n x + b_n\big)\Big| \leq \max\Big(\Lambda(x), \Phi^n\big(a_n x + b_n\big)\Big) < \frac{1}{2^n} \tag{32}$$

where we have used that for $n \geq 3$,

$$\exp\left\{-e^{b_n^2+1}\right\} \leq \frac{1}{2^n}.$$

It is easy to see that this inequality holds for $n \leq 5$. To prove it for $n \geq 6$, notice that it is equivalent to see that

$$e^{b_n^2+1} \geq (\log 2)n.$$

Now, by Proposition 4,

$$e^{b_n^2+1} \geq e^{1+2\log n - \log(4\pi \log n)} = \frac{e\, n^2}{4\pi \log n}.$$

Hence, it suffices to prove that, for $n \geq 6$, $e\, n^2/(4\pi \log n) \geq (\log 2)n$ and this result follows studying the function $y/\log y$ and its derivatives.

Finally, since for $y > 1$ the function $(1/2)^y \log y$ is decreasing, inequality (32) implies that

$$\Big|\Lambda(x) - \Phi^n\big(a_n x + b_n\big)\Big| < \frac{\log n_0}{2^{n_0}} \frac{1}{\log n}. \tag{33}$$

2. Case $x \in [-b_n/a_n, -1.25 \log n]$. As in the first case we will bound separately $\Lambda(x)$ and $\Phi^n(a_n x + b_n)$.

We start studying $\Lambda(x)$. We get that

$$0 \leq \Lambda(x) \leq \Lambda(-1.25 \log b_n) < \Lambda(-\log b_n) \leq \frac{4}{e^2} \frac{1}{b_n^2} \leq \frac{4 \log n_0}{e^2 b_{n_0}^2} \frac{1}{\log n}. \tag{34}$$

where we have used Proposition 5 and that

$$\Lambda(-y) \leq \frac{4}{e^2}\, e^{-2y} \tag{35}$$

for all $y \geq 0$. Notice that this inequality holds because $\max_{z \geq 0}\left(z^2 e^{-z}\right) = 4/e^2$.

Let us consider now the term $\Phi^n(a_n x + b_n)$. Recall that by (23),

$$\log\, \Phi^n(a_n x + b_n) = -\exp\Big\{-\int_{b_n}^{a_n x + b_n} V(t)\, dt\Big\} - n S_n(x), \tag{36}$$

with $S_n(x) > 0$. Hence,

$$\Phi^n(a_n x + b_n) \leq e^{-\exp\left\{\int_{a_n x + b_n}^{b_n} V(t)\, dt\right\}}. \tag{37}$$

By (19) we get

$$\int_{a_n x + b_n}^{b_n} V(t)\, dt \geq \int_{a_n x + b_n}^{b_n} t\, dt = -a_n b_n x - a_n^2 \frac{x^2}{2} := g(x).$$

Notice that on the interval $[-b_n/a_n, -1.25\log b_n]$ the function g is decreasing and then for all x in this interval, $g(x) \geq g(-1.25\log b_n)$. Thus,

$$\begin{aligned}
\int_{a_n x+b_n}^{b_n} V(t)\,dt &\geq \frac{5}{4}a_n b_n \log b_n - \frac{25}{32}a_n^2(\log b_n)^2 \\
&= \frac{5}{4}\frac{b_n^2}{b_n^2+1}\log b_n - \frac{25b_n^2}{32(b_n^2+1)^2}(\log b_n)^2 \\
&\geq \frac{5}{4}\frac{b_n^2}{b_n^2+1}\log b_n - \max_{y>1}\left(\frac{25y^2}{32(y^2+1)^2}\log y\right)\log b_n \\
&\geq \frac{5}{4}\frac{b_n^2}{b_n^2+1}\log b_n - \frac{1}{10}\log b_n = \frac{23b_n^2-2}{20(b_n^2+1)}\log b_n, \qquad (38)
\end{aligned}$$

where we have used that

$$\max_{y\geq 1}\left(\frac{25y^2}{32(y^2+1)^2}\log y\right) < \frac{1}{10}, \qquad (39)$$

and that $b_n > 1$ for $n \geq 16$. The above inequality follows studying the function $h(y) = (y^2\log y)/(y^2+1)^2$. In fact,

$$h'(y) = y\frac{y^2+1+2(1-y^2)\log y}{(y^2+1)^3},$$

and its sign, when $y > 1$, is the contrary of the sign of

$$H(y) = \log y - \frac{y^2+1}{2(y^2-1)},$$

which can be easily studied because $H'(y) = (y^4+1)/(y(y^2-1)^2) > 0$. Hence, for $y > 1$, the function h is positive and increasing until some value $y = y^*$ and then decreases monotonically towards zero. By Bolzano's Theorem, $y^* \in (\underline{y}, \overline{y}) := (2.16, 2.17)$. Hence

$$\max_{y>1} h(y) < \frac{\overline{y}^{\,2}\log\overline{y}}{(\underline{y}^2+1)^2} \quad \text{and} \quad \frac{25}{32}\frac{\overline{y}^{\,2}\log\overline{y}}{(\underline{y}^2+1)^2} < \frac{1}{10}.$$

Combining (37), (38) and (35) we obtain that

$$\Phi^n(a_n x+b_n) \leq \Lambda\left(-\int_{a_n x+b_n}^{b_n} V(t)\,dt\right) \leq \frac{4}{e^2}\exp\left\{\frac{2-23b_n^2}{10(b_n^2+1)}\log b_n\right\}.$$

Hence, once we prove that

$$\max_{y\geq 1} P(y) < \frac{2}{3}, \qquad (40)$$

where

$$P(y) = \frac{4}{e^2}y^2\exp\left\{\frac{2-23y^2}{10(y^2+1)}\log y\right\} = \frac{4}{e^2}y^2 Q(y),$$

we will have that

$$\Phi^n(a_n x+b_n) \leq \frac{2}{3b_n^2} \leq \frac{2\log n_0}{3b_{n_0}^2}\frac{1}{\log n}.$$

where note that we have used once more Proposition 5.

Joining (34) and the above inequality we get that when $x \in [-b_n/a_n, -1.25\log n]$,

$$\left|\Phi^n(a_n x+b_n) - \Lambda(x)\right| \leq \frac{2\log n_0}{3b_{n_0}^2}\frac{1}{\log n}, \qquad (41)$$

Hence, to end this part of the proof we need to prove (40). To study the function $P(y)$ we compute

$$P'(y) = \frac{2}{5e^2}\frac{y}{(1+y^2)^2}\,Q(y)\,q(y), \quad \text{with} \quad q(y) = (22-3y^2)(1+y^2) - 50y^2\log y.$$

Moreover, for $y \geq 1$, $q'(y) = -100\,y\log y - 12(1+y^2)y < 0$. Joining all this information we get that for $y \geq 1$, the function $P'(y)$ is decreasing and has a unique zero y^* and, by Bolzano's Theorem, $y^* \in (\underline{y}, \overline{y}) := (1.532, 1.533)$. Therefore the function P is increasing in $[1, y^*)$ and decreasing in (y^*, ∞). As a consequence,

$$\max_{y\geq 1} P(y) = P(y^*) < \frac{4}{e^2}\overline{y}^2 \exp\left\{\frac{2-23\underline{y}^2}{10(\overline{y}^2+1)}\log\underline{y}\right\} < 0.66 < \frac{2}{3},$$

as we wanted to prove.

3. Case $x \in (-1.25\log n, 0)$. Using Proposition 6 we write

$$\Phi^n(a_nx+b_n) - \Lambda(x) = e^{-nS_n(x)}\Lambda(x)\Big(\frac{\Lambda\left(I_n(x)\right)}{\Lambda(x)} - 1\Big) + \Lambda(x)\left(e^{-nS_n(x)} - 1\right).$$

Hence

$$|\Phi^n(a_nx+b_n) - \Lambda(x)| \leq \Lambda(x)\Big|\frac{\Lambda\left(I_n(x)\right)}{\Lambda(x)} - 1\Big| + \Lambda(x)\left|e^{-nS_n(x)} - 1\right|. \tag{42}$$

We start proving that

$$-I_n(x) = \int_{a_nx+b_n}^{b_n} V(t)\,dt \leq -x. \tag{43}$$

Recall that by (19), for $t > 0$, $V(t) \leq t + 1/t$ and moreover that V is an increasing function. Hence

$$\int_{a_nx+b_n}^{b_n} V(t)\,dt \leq -V(b_n)a_nx \leq -\Big(b_n + \frac{1}{b_n}\Big)a_nx = -\frac{b_n^2+1}{b_n}a_nx = -x.$$

Therefore, for the first term of the right hand side of (42) we have

$$\begin{aligned}\Lambda(x)\Big|\frac{\Lambda\left(I_n(x)\right)}{\Lambda(x)} - 1\Big| &= \Lambda(x)\left|\exp\left(e^{-x} - e^{-I_n(x)}\right) - 1\right| \\ &= \Lambda(x)\left\{\exp\left(e^{-x} - e^{-I_n(x)}\right) - 1\right\} \\ &\leq \Lambda(x)\exp\left\{e^{-x} - e^{-I_n(x)}\right\}\left(e^{-x} - e^{-I_n(x)}\right),\end{aligned} \tag{44}$$

where in the last inequality we have applied the Mean Value Theorem to e^x.

Now, notice that taking into account that for $y \geq 0$, $1 - e^{-y} \leq y$,

$$\begin{aligned}e^{-x} - e^{-I_n(x)} &= e^{-x}\left(1 - e^{x+\int_{a_nx+b_n}^{b_n} V(t)\,dt}\right) \leq e^{-x}\left(-x - \int_{a_nx+b_n}^{b_n} V(t)\,dt\right) \\ &\leq e^{-x}\left((a_nb_n - 1)x - \frac{a_n^2}{2}x^2\right) \leq e^{-x}\Big(-x + \frac{x^2}{2}\Big)\frac{1}{b_n^2},\end{aligned} \tag{45}$$

where in the last inequalities we have used first that $V(t) > t$, and later that a_n^2 and $1 - a_nb_n$ are both smaller than $1/b_n^2$.

To continue, notice that since $-x \leq 1.25\log b_n$, then $b_n^2 \geq \exp(-8x/5)$. Hence, from (44) and (45),

$$\Lambda(x)\Big|\frac{\Lambda\left(I_n(x)\right)}{\Lambda(x)} - 1\Big| \leq Q(x)\frac{1}{b_n^2},$$

where

$$Q(x) = \Big(- x + \frac{x^2}{2}\Big) \exp\Big\{ -x - e^{-x} + \Big(- x + \frac{x^2}{2}\Big) e^{3x/5}\Big\} = \Big(- x + \frac{x^2}{2}\Big) T(x).$$

We claim

$$0 < \max_{x<0} Q(x) < 0.63. \tag{46}$$

Therefore

$$\Lambda(x) \Big| \frac{\Lambda\left(I_n(x)\right)}{\Lambda(x)} - 1 \Big| \leq \frac{0.63}{b_n^2}. \tag{47}$$

Let us prove now the inequality (46) given in the above claim. Notice that

$$Q'(x) = \frac{1}{20} T(x)\, t(x) = \frac{1}{20} T(x) \left(t_0(x) + t_1(x) e^{-x} + t_2(x) e^{3x/5} \right),$$

where $t_0(x) = -10(x^2 - 4x + 2)$, $t_1(x) = 10x(x-2)$ and $t_2(x) = x(x-2)(3x^2 + 4x - 10)$. To study the sign of $Q'(x)$ we consider the function $t(x)$ for $x < 0$. We get that

$$t''(x) = -20 + s_1(x) e^{-x} + s_2(x) e^{3x/5},$$

where $s_1(x) = 10(x^2 - 6x + 6) > 0$ and $s_2(x) = (27x^4 + 342x^3 + 558x^2 - 1200x - 300)/25$. It is clear that

$$t''(x) \geq -20 + s_1(x) + s_2(x) e^{3x/5} = s_0(x) + s_2(x) e^{3x/5} = S(x),$$

with $s_0(x) = 10(x^2 - 6x + 4)$. Hence, if we prove that $S(x) > 0$ we will have the convexity of $t(x)$. Joining this information with the fact that $t(0) = -20 < 0, t'(0) = 40 > 0$ and that $t(x)$ tends to infinity when x goes to minus infinity, we obtain that $t(x)$ has a unique zero $x = x^*$ in $(-\infty, 0)$. By Bolzano's Theorem $x^* \in (\underline{x}, \overline{x}) := (-1.051, -1.050)$. Finally Q is increasing on $(-\infty, x^*)$ and decreasing in $(x^*, 0)$ and therefore,

$$\max_{x<0} Q(x) = Q(x^*) < \Big(- \underline{x} + \frac{\underline{x}^2}{2}\Big) \exp\Big\{ -\underline{x} - e^{-\overline{x}} + \Big(- \underline{x} + \frac{\underline{x}^2}{2}\Big) e^{3\overline{x}/5}\Big\} < 0.63,$$

as we wanted to prove. That $S(x) > 0$ for $x < 0$, can be proved by using similar arguments and we omit the details.

To end the proof it remains to study the term $\Lambda(x)|e^{-nS_n(x)} - 1|$, where recall that from Proposition 6,

$$S_n(x) \leq \frac{C_n^2(x)}{2(1 - C_n(x))} \quad \text{and} \quad C_n(x) = \frac{1}{n} \exp\Big\{ \int_{a_n x + b_n}^{b_n} V(t)\, dt \Big\} = 1 - \Phi(a_n x + b_n).$$

Hence

$$nS_n(x) \leq \frac{\left(\exp\Big\{ \int_{a_n x + b_n}^{b_n} V(t)\, dt \Big\} \right)^2}{2n\Phi(a_n x + b_n)} \leq \frac{1}{n} e^{-2x},$$

where we have used (43) and that $\Phi(a_n x + b_n) > 1/2$, because $a_n x + b_n > 0$.

Using this inequality, that $1 - e^{-y} \leq y$ for $y \geq 0$ and that $e^{-x} \geq 1 - x + x^2/2$, for $x \leq 0$, we obtain

$$\begin{aligned}
\Lambda(x) \big| e^{-nS_n(x)} - 1 \big| &\leq \frac{1}{n} \Lambda(x) e^{-2x} \leq \frac{1}{n} e^{-1 + x - x^2/2} e^{-2x} \\
&\leq \frac{1}{n} e^{-1/2} e^{-(x+1)^2/2} \leq \frac{\log n}{\sqrt{e} n} \frac{1}{\log n} \leq \frac{\log n_0}{\sqrt{e} n_0} \frac{1}{\log n}.
\end{aligned} \tag{48}$$

Joining (47) and (48) and using Proposition 5 we arrive to

$$\big| \Phi^n(a_n x + b_n) - \Lambda(x) \big| \leq \left(\frac{0.63 \log n_0}{b_{n_0}^2} + \frac{\log n_0}{\sqrt{e} n_0} \right) \frac{1}{\log n}, \tag{49}$$

for the values of x considered in this case.

Finally, collecting the right hand terms of inequalities (33), (41) and (49) we have that for $n \geq 16$,

$$\max\left(\frac{\log n_0}{2^{n_0}}, \frac{2\log n_0}{3b_{n_0}^2}, \Big(\frac{0.63\log n_0}{b_{n_0}^2} + \frac{\log n_0}{\sqrt{e}n_0}\Big)\right) < \frac{2\log n_0}{3b_{n_0}^2} + \frac{\log n_0}{\sqrt{e}n_0}.$$

Hence

$$\left|\Phi^n(a_n x + b_n) - \Lambda(x)\right| < \Big(\frac{2\log n_0}{3b_{n_0}^2} + \frac{\log n_0}{\sqrt{e}n_0}\Big)\frac{1}{\log n},$$

and the proposition follows. □

4.3 Global rate of convergence: proof of Theorem 1

The first part of the Theorem 1 is a straightforward consequence of Propositions 7 and 8: For $n_0 \geq 16$, because it is easy to prove that $C^-(n_0) > C^+(n_0)$. For $5 \leq n_0 \leq 15$ because $C^-(n_0) = 1$ and $C^+(n_0) < 1$.

Proposition 4 provides upper and lower bounds for b_n^2. These bounds substituted in $C(n_0)$ give easily that $\lim_{n_0\to\infty} C(n_0) = 1/3$. □

4.4 Other norming constants a_n

If instead of $a_n^\circ = A_{\mathcal{F}}(b_n) = b_n/(1+b_n^2)$ it is used $A_{\mathcal{H}}(b_n) = 1/b_n$, applying similar tools that in the proof of Proposition 7 we get the following result:

Proposition 9. *Given $n_0 \geq 2$, for all $n \geq n_0$ it holds that*

$$\sup_{x\geq 0} \left|\Phi^n\big(A_{\mathcal{H}}(b_n)\,x + b_n\big) - \Lambda(x)\right| < \frac{\overline{C}^+(n_0)}{\log n},$$

with

$$\overline{C}^+(n_0) = \frac{\sqrt{2}+1}{e^{\sqrt{2}}}\frac{\log n_0}{b_{n_0}^2} + \frac{\log n_0}{2(n_0-1)}.$$

Proof. Starting as in the proof of Proposition 7 we obtain that

$$\left|\Phi^n(A_{\mathcal{H}}(b_n)x + b_n) - \Lambda(x)\right| \leq \left|\Lambda\left(I_n(x)\right) - \Lambda(x)\right| + \frac{\log n_0}{2(n_0-1)}\frac{1}{\log n}, \tag{50}$$

where here

$$I_n(x) = \int_{b_n}^{A_{\mathcal{H}}(b_n)x+b_n} V(t)\,dt, \quad \text{and} \quad A_{\mathcal{H}}(b_n) = \frac{1}{b_n}.$$

To study the remainder left hand term of (50) let us prove first that under our hypotheses $I_n(x) > x$. Notice that since $A_{\mathcal{C}}(t) < 1/t$, $A_{\mathcal{C}}(b_n) < A_{\mathcal{H}}(b_n)$. Then, by the Mean Value Theorem, there is $x_1 \in [0, x]$ such that

$$\begin{aligned} I_n(x) = \int_{b_n}^{A_{\mathcal{H}}(b_n)x+b_n} V(t)\,dt &> \int_{b_n}^{A_{\mathcal{C}}(b_n)x+b_n} V(t)\,dt \\ &= V\big(A_{\mathcal{C}}(b_n)x_1 + b_n\big)A_{\mathcal{C}}(b_n)\,x = \frac{A_{\mathcal{C}}(b_n)}{A_{\mathcal{C}}\big(A_{\mathcal{C}}(b_n)x_1 + b_n\big)}\,x > x, \end{aligned}$$

where in the last step we have used that $A_{\mathcal{C}}$ is decreasing. Then

$$\begin{aligned} \left|\Lambda\left(I_n(x)\right) - \Lambda(x)\right| &= \Lambda\left(I_n(x)\right) - \Lambda(x) \leq \Lambda'(x)\big(I_n(x) - x\big) \\ &= \Lambda(x)e^{-x}\big(I_n(x) - x\big) \leq e^{-x}\big(I_n(x) - x\big), \end{aligned} \tag{51}$$

using once more that for $x > 0$, $\Lambda(x)$ is increasing, and the Mean Value Theorem.

Now, taking into account that $V(t) \leq t + 1/t$ (see (19)) and again that for $y > -1$, $\log(1+y) \leq y$, we obtain that

$$I_n(x) - x = \int_{b_n}^{x/b_n+b_n} V(t)\,dt - x \leq \frac{1}{2}\frac{x^2}{b_n^2} + \log\left(\frac{x}{b_n^2}+1\right) < \frac{x^2/2+x}{b_n^2}. \tag{52}$$

Moreover, the following bound is immediate: for $x \geq 0$,

$$0 \leq e^{-x}(x^2/2+x) \leq e^{-\sqrt{2}}(\sqrt{2}+1). \tag{53}$$

Joining (51), (52) and (53), we get

$$\Big|\Lambda\left(I_n(x)\right) - \Lambda(x)\Big| \leq \frac{\sqrt{2}+1}{e^{\sqrt{2}}b_n^2}. \tag{54}$$

Finally, applying Proposition 5 the result follows. □

Remark 10. (i) Notice that when $x \geq 0$, applying Proposition 7 and Proposition 4 we have that choosing $a_n^\circ = A_{\mathcal{F}}(b_n) = b_n/(1+b_n^2)$ we get that

$$\lim_{n_0\to\infty} C^+(n_0) = \frac{1}{2e} \approx 0.184,$$

while Proposition 9 implies that when $a_n = A_{\mathcal{H}}(b_n) = 1/b_n$ then

$$\lim_{n_0\to\infty} \overline{C}^+(n_0) = \frac{\sqrt{2}+1}{2e^{\sqrt{2}}} \approx 0.294$$

The above results are coherent with the numerical results presented in Table 2 and show that the first choice $a_n^\circ = A_{\mathcal{F}}(b_n)$ gives best approximations. We do not develop here the case $x < 0$ for $A_{\mathcal{H}}(b_n)$.

(ii) It is worth noting that since $A_{\mathcal{C}}(t) < 1/t$, $A_{\mathcal{C}}(b_n) < A_{\mathcal{H}}(b_n)$, it is easy to see that Proposition 9 also holds replacing $A_{\mathcal{H}}(b_n)$ by $A_{\mathcal{C}}(b_n)$. Nevertheless the bound given by this result seems less accurate than the ones provided in Propositions 7 and 9, see again Table 2.

5 Explicit norming constants

Since the expression b_n^* is not explicit, by using asymptotic analysis it can be deduced expressions asymptotically equivalent for b_n^* and a_n^* (they satisfy Property 2)

$$\beta_n^* = (2\log n)^{1/2} - \frac{\log\log n + \log(4\pi)}{2\,(2\log n)^{1/2}} \tag{55}$$

and

$$\alpha_n^* = (2\log n)^{-1/2}.$$

(see, for example, Resnick [9, pp. 71–72]). An easy way to deduce these constants and to suggest other ones more suited to previous results is to use the Lambert W function and its extensions.

5.1 Lambert W function and extensions

For $t > 0$, the equation $ye^y = t$ has a unique real positive solution y, which determines (for $t > 0$) the principal branch of the real Lambert W function, that means, $W(t)$ satisfies

$$W(t)\,e^{W(t)} = t,$$

and $\lim_{t\to\infty} W(t) = \infty$ (see Corless *et al.* [3] for a complete overview of Lambert W function and many applications). The asymptotic expansion of this function is given by Corless *et al.* [3, pp. 22 and 23], see also De Bruijn [4, pp. 25–27].

$$W(t) = \log t - \log\log t + \frac{\log\log t}{\log t} + O\Big(\Big(\frac{\log\log t}{\log t}\Big)^2\Big), \quad t \to \infty. \tag{56}$$

For $\gamma \neq 0$, Comtet [2] extended that expansion to the (unique) positive solution y of the equation

$$y^\gamma e^y = t$$

such that $y \to \infty$ when $t \to \infty$. Later, Robin [10] and Salvy [11] extended Comtet [2] results in order to deduce an asymptotic expansion of the solution of the equation

$$y^\gamma e^y D\Big(\frac{1}{y}\Big) = t, \tag{57}$$

where

$$D(y) = \sum_{n=0}^{\infty} d_n y^n, \text{ with } d_0 \neq 0,$$

is a power series convergent in a neighborhood of the origin. Denote by $U_D(t)$ that solution. We are only interested on the case $\gamma = 1$ and $d_0 = 1$, and for this case, Robin [10] and Salvy [11] prove

$$U_D(t) = \log t - \log\log t + \frac{\log\log t - d_1}{\log t} + \frac{Q_2(\log\log t)}{(\log t)^2} + o\Big(\frac{1}{(\log t)^2}\Big), \ t \to \infty,$$

where Q_2 is a polynomial of degree 2, whose coefficients depend on D. The above expression implies that

$$U_D(t) = \log t - \log\log t + \frac{\log\log t - d_1}{\log t} + O\Big(\Big(\frac{\log\log t}{\log t}\Big)^2\Big), \ t \to \infty. \tag{58}$$

5.2 Return to the norming constants

Thanks to (11), the constant b_n^* can be written in terms of the principal branch of Lambert function:

$$b_n^* = \Big(W\big(n^2/(2\pi)\big)\Big)^{1/2}.$$

Hence, from, (56),

$$\begin{aligned} b_n^* &= \Big(\log\big(n^2/(2\pi)\big) - \log\log\big(n^2/(2\pi)\big) + O\Big(\frac{\log\log n}{\log n}\Big)\Big)^{1/2} \\ &= (2\log n)^{1/2} - \frac{\log(4\pi\log n)}{2\,(2\log n)^{1/2}} + O\Big(\frac{(\log\log n)^2}{(\log n)^{3/2}}\Big) \\ &= \beta_n^* + O\Big(\frac{(\log\log n)^2}{(\log n)^{3/2}}\Big), \quad n \to \infty. \end{aligned}$$

Notice that if we introduce the following sequence

$$\overline{\beta}_n^* = \Big(\log\big(n^2/(2\pi)\big) - \log\log\big(n^2/(2\pi)\big) + \frac{\log\log(n^2/(2\pi))}{\log(n^2/(2\pi))}\Big)^{1/2},$$

we obtain a better approximation to b_n^* because

$$b_n^* = \overline{\beta}_n^* + O\left(\frac{(\log\log n)^2}{(\log n)^{5/2}}\right), \quad n \to \infty.$$

In any case, Theorem 1 suggests that the utilization of an approximation to b_n rather than an approximation to b_n^* likely will provide more velocity of convergence. To this end, in next proposition we compute an asymptotic expansion of b_n at infinity using the bounds (17) for the Mills ratio and the function U_D introduced in (58).

Proposition 11. *It holds that*

$$b_n = \overline{\beta}_n + O\left(\frac{(\log\log n)^2}{(\log n)^{5/2}}\right), \quad n \to \infty, \tag{59}$$

where

$$\overline{\beta}_n = \left(\log\left(n^2/(2\pi)\right) - \log\log\left(n^2/(2\pi)\right) + \frac{\log\log(n^2/(2\pi)) - 2}{\log(n^2/(2\pi))}\right)^{1/2}.$$

Proof. By inequalities (17) we know that for $x > 0$,

$$r(x)\phi(x) < 1 - \Phi(x) < R(x)\phi(x),$$

where

$$r(x) = \frac{x}{x^2+1} \quad \text{and} \quad R(x) = \frac{x^2+2}{x^3+3x}.$$

For n large enough, let v_n (resp. V_n) be the solution of the equation $r(x)\phi(x) = 1/n$ (resp. $R(x)\phi(x) = 1/n$). Recall that b_n satisfies $1 - \Phi(b_n) = 1/n$. Therefore, for these values of n it holds that

$$v_n \le b_n \le V_n. \tag{60}$$

Let us compute the asymptotic expansions at infinity of $\{v_n\}$ and $\{V_n\}$. Notice that v_n satisfies the equation

$$\frac{n^2}{2\pi} = x^2 e^{x^2}\Big(1 + \frac{1}{x^2}\Big)^2 = x^2 e^{x^2}\Big(1 + 2\frac{1}{x^2} + \frac{1}{x^4}\Big),$$

while V_n satisfies

$$\frac{n^2}{2\pi} = x^2 e^{x^2}\left(\frac{1+\frac{3}{x^2}}{1+\frac{2}{x^2}}\right)^2 = x^2 e^{x^2}\left(1 + 2\frac{1}{x^2} + O\Big(\frac{1}{x^4}\Big)\right).$$

By changing x^2 by y, and with the notations of Subsection 5.1, both v_n and V_n, are $\big(U_D\big(n^2/(2\pi)\big)\big)^{1/2}$, for some analytic functions D, both satisfying

$$D(y) = 1 + 2y + O\big(y^2\big).$$

Then, from (58) and (60), we arrive to the same asymptotic expansions for v_n, b_n and V_n. Specifically,

$$b_n = \left(\log\left(n^2/(2\pi)\right) - \log\log\left(n^2/(2\pi)\right) + \frac{\log\log\left(n^2/(2\pi)\right) - 2}{\log\left(n^2/(2\pi)\right)} + O\Big(\Big(\frac{\log\log n}{\log n}\Big)^2\Big)\right)^{1/2}.$$

From the above expression, (59) follows from $\sqrt{w+a} - \sqrt{w} = a/(\sqrt{w+a} + \sqrt{w})$. $\square$

Remark 12. Using the same tools that in the proof of the above proposition we obtain that

(i) $b_n - b_n^* = O\left(\dfrac{1}{(\log n)^{3/2}}\right), \quad n \to \infty,$

(ii) $\overline{\beta}_n - \beta_n^* = O\left(\dfrac{(\log\log n)^2}{(\log n)^{3/2}}\right), \quad n \to \infty,$

(iii) $\overline{\beta}_n - \overline{\beta}_n^* = O\left(\dfrac{1}{(\log n)^{3/2}}\right), \quad n \to \infty.$

Remark 13. It is also possible to construct some approximations of b_n^* and b_n that are improvements of (55) adding some suitable terms. In fact, if we define:

$$\widetilde{\beta}_n^* = \sqrt{2\log n} - \frac{\log(4\pi\log n)}{2\sqrt{2\log n}} - \frac{\big(\log(4\pi\log n)\big)^2 - 4\log(4\pi\log n)}{8\sqrt{(2\log n)^3}},$$

$$\widetilde{\beta}_n = \sqrt{2\log n} - \frac{\log(4\pi\log n)}{2\sqrt{2\log n}} - \frac{\big(\log(4\pi\log n)\big)^2 - 4\log(4\pi\log n) + 8}{8\sqrt{(2\log n)^3}},$$

it also holds that

$$b_n^* = \widetilde{\beta}_n^* + O\left(\frac{(\log\log n)^2}{(\log n)^{5/2}}\right) \quad \text{and} \quad b_n = \widetilde{\beta}_n + O\left(\frac{(\log\log n)^2}{(\log n)^{5/2}}\right) \qquad n \to \infty.$$

Nevertheless the approximations $\overline{\beta}_n^*$ and $\overline{\beta}_n$, respectively, are sharper, specially for small n.

5.3 From $\overline{\beta}_n$ to β_n

As we have seen in the previous subsection, $\overline{\beta}_n$ is a very good approximation of b_n. Nevertheless, for each $p, q \in \mathbb{R}$, if we introduce the new constants

$$B_n(p,q) = \left(\log\big(n^2/(2\pi)\big) - \log\log\big(n^2/(2\pi)\big) + \frac{\log\big(\log(n^2) + p\big) - 2}{\log\big(n^2\big) + q}\right)^{1/2},$$

it also holds that

$$b_n = B_n(p,q) + O\left(\frac{(\log\log n)^2}{(\log n)^{5/2}}\right), \quad n \to \infty.$$

In particular, $\overline{\beta}_n = B_n\big(-\log(2\pi), -\log(2\pi)\big)$.

To obtain some values of p and q that provide better approximations to b_n, at least for n in the most used range $[10, 10^5]$, we proceed as follows: For simplicity we fix $q = -\log(2\pi)$ and consider p as a free parameter to be determined. For a given $m \in \mathbb{N}$, we consider the set of $m-9$ equations

$$b_k - B_k\big(p, -\log(2\pi)\big) = 0, \quad k = 10, 11, \ldots, m.$$

The actual values b_k are obtained numerically. For each k, let p_k be the solution of the corresponding equation, which is also obtained numerically. Then we define

$$\widehat{p}(m) = \frac{1}{m-9}\sum_{k=10}^{m} p_k.$$

Notice that $\widehat{p}(m)$ can be interpreted as the "best" solution for the incompatible system formed by the corresponding $m-9$ equations. We have obtained that $\widehat{p}(10^2) \approx 0.59$, $\widehat{p}(10^3) \approx 0.47$, $\widehat{p}(10^4) \approx 0.47$, and $\widehat{p}(10^5) \approx 0.52$. These values suggest to consider $p = 1/2$ as a candidate to have an approximation of b_n that is good both for $n \in [10, 10^5]$ and for n large enough. In short we consider

$$\beta_n = B_n\big(1/2, -\log(2\pi)\big),$$

that is precisely the expression (6) given in the introduction.

In table 3 there is a numerical comparison between all the constants involved in this section for different sample size. These results illustrate that the suggested new constant β_n is a very good approximation for b_n, and that it is sharper than $\overline{\beta}_n$, specially for small values of n. Also $\overline{\beta}_n^*$ is a good approximation to b_n^*, whereas β_n^* approximates b_n^*, but more slowly. The computations to get the table are done with Maple.

n	10	10^2	10^5	10^{10}	10^{30}	10^{60}
b_n	1.28155	2.32635	4.26489	6.36134	11.46402	16.39728
β_n	1.27115	2.32632	4.26488	6.36132	11.46402	16.39728
$\overline{\beta}_n$	1.18090	2.31828	4.26430	6.36123	11.46401	16.39728
b_n^*	1.43165	2.37533	4.27575	6.36492	11.46467	16.39750
$\overline{\beta}_n^*$	1.45508	2.37607	4.27535	6.36478	11.46465	16.39750
β_n^*	1.36192	2.36625	4.28019	6.36855	11.46611	16.39821

Table 3. Comparison of the standard constants β_n^* and the constants $\overline{\beta}_n^*$ with b_n^* and the proposed constants $\overline{\beta}_n$ and β_n with b_n.

Conclusions

As a corollary of Theorem 1, Proposition 11 and the computations of this section, we propose (6),

$$\beta_n = \left(\log\big(n^2/(2\pi)\big) - \log\log\big(n^2/(2\pi)\big) + \frac{\log\big(\log(n^2)+1/2\big)-2}{\log\big(n^2/(2\pi)\big)} \right)^{1/2},$$

that is a very good approximation for b_n, as one of the norming constants for the maximum of n i.i.d. standard normal random variables. Also, in agreement with Remark 10 and Table 2, it seems also convenient to utilize always $A_{\mathcal{F}}$. So we propose, instead of α_n^* to use, together with b_n, the norming constant

$$\alpha_n = \frac{\beta_n}{1+\beta_n^2}.$$

Acknowledgments

The first author was partially supported by grants MINECO/FEDER reference MTM2008-03437 and Generalitat de Catalunya reference 2009-SGR410. The second and third authors by grants MINECO/FEDER reference MTM2009-08869 and MINECO reference MTM2012-33937

References

[1] BARICZ, Á., Mills' ratio: monotonicity patterns and functional inequalities. *J. Math. Anal. Appl.* **340** (2008) 1362–1370.

[2] COMTET, L., Inversion de $y^\alpha e^y$ et $y \log^\alpha y$ au moyen des nombres de Stirling, *C. R. Acad. Sc. Paris*, Serie A, t. **270** (1970) 1085–1088.

[3] CORLESS, R.M., GONNET, G.H., HARE, D.E.G., JEFFREY, D.J., KNUTH, D.E., On the Lambert W function. *Adv. Comput. Math.* **5** (1996) 329–359.

[4] DE BRUIJN, N.G., *Asymptotic Methods in Analysis*, Dover, New York, 1981.

[5] FISHER, R.A. AND TIPPETT, L.H.C., Limiting forms of the frequency distribution of the largest or smallest member of a sample. *Proc. Cambridge Philos. Soc.* **24** (1928) 180–190.

[6] GASULL, A. AND UTZET, F., Approximating Mills ratio. Arxiv 1307.3433 (2013)

[7] HALL, P., On the rate of convergence of normal extremes. *J. Appl. Probab.* **16** (1979) 433–439.

[8] PINELIS, I., Monotonicity properties of the relative error of a Padé approximation for Mills' ratio. *J. Inequal. Pure Appl. Math.* **3** (2002), Article 20, 8 pp.

[9] RESNICK, S.I., *Extreme Values, Regular Variation, and Point Processes*, Springer, Berlin, 1987.

[10] ROBIN, G., Permanence de relations de recurrence dans certains developpements asymptotiques. *Publications de l'Institut Mathématique*, Nouv. sér., tome **43** (57) (1988) 17–25.

[11] SALVY, B., Fast computation of some asymptotic functional inverses. *J. Symbolic Comput.* **17** (1994) 227–236.

[12] SMALL, C.G., *Expansions and Asymptotics for Statistics*, Chapman and Hall/CRC Press, Boca Raton, 2010.