On norming constants for normal maxima

Joint work with A. Gasull and F. Utzet

Universitat Autònoma de Barcelona

Barcelona, May 2015

Outline

- Basic results on Extreme Value Theory.
- Hall's result on the velocity of convergence of the normal maxima.
- Our main theorem and some ideas about its proof.
- Explicit expressions for the norming constants

Outline

- Basic results on Extreme Value Theory.
- Hall's result on the velocity of convergence of the normal maxima.
- Our main theorem and some ideas about its proof.
- Explicit expressions for the norming constants

Outline

- Basic results on Extreme Value Theory.
- Hall's result on the velocity of convergence of the normal maxima.
- Our main theorem and some ideas about its proof.
- Explicit expressions for the norming constants

Outline

- Basic results on Extreme Value Theory.
- Hall's result on the velocity of convergence of the normal maxima.
- Our main theorem and some ideas about its proof.
- Explicit expressions for the norming constants

Problem:

Given $X_1, , X_2, , \ldots$ sequence of i. i. d. random variables with distribution function F, find the limit behavior of

$$M_n = \max(X_1, \ldots, X_n).$$

Concretely, we are interested in the case that the X_n have standard normal distribution.

We first give some some important results on Extreme value Theory.
A first easy result is:

Defining

$$x_F = \sup\{x \in \mathbb{R} : F(x) < 1\} \leq \infty$$

as the right endpoint of F, then

$$M_n \longrightarrow x_F \quad \text{a. s..}$$

Problem:
Given $X_1, , X_2, , \ldots$ sequence of i. i. d. random variables with distribution function F, find the limit behavior of

$$M_n = \max(X_1, \ldots, X_n).$$

Concretely, we are interested in the case that the X_n have standard normal distribution.

We first give some some important results on Extreme value Theory.
A first easy result is:

Defining

$$x_F = \sup\{x \in \mathbb{R} : F(x) < 1\} \leq \infty$$

as the right endpoint of F, then

$$M_n \longrightarrow x_F \quad \text{a. s..}$$

Indeed, for any $\omega \in \Omega$, $M_n(\omega)$ is an increasing sequence, so the limit

$$M(\omega) := \lim_n M_n(\omega)$$

exists (it can be $+\infty$).

In order to identify the limit, first consider the case $x_F < \infty$. We have, for any $x < x_F$

$$P\{M_n \le x\} = \big(F(x)\big)^n \to 0,$$

so, M_n converges in law to x_F (observe that $P\{M_n \le x_F\} = 1$).

In the case $x_F = \infty$, we have that

$$P\{M = \infty\} = \lim_{K \to \infty} P\{M > K\}.$$

But,

$$P\{M > K\} \ge P\{M_n > K\} = 1 - F(K)^n \longrightarrow 1, \text{ as } n \to \infty.$$

Indeed, for any $\omega \in \Omega$, $M_n(\omega)$ is an increasing sequence, so the limit

$$M(\omega) := \lim_n M_n(\omega)$$

exists (it can be $+\infty$).
In order to identify the limit, first consider the case $x_F < \infty$. We have, for any $x < x_F$

$$P\{M_n \leq x\} = \big(F(x)\big)^n \to 0,$$

so, M_n converges in law to x_F (observe that $P\{M_n \leq x_F\} = 1$).
In the case $x_F = \infty$, we have that

$$P\{M = \infty\} = \lim_{K \to \infty} P\{M > K\}.$$

But,

$$P\{M > K\} \geq P\{M_n > K\} = 1 - F(K)^n \longrightarrow 1, \text{ as } n \to \infty.$$

Indeed, for any $\omega \in \Omega$, $M_n(\omega)$ is an increasing sequence, so the limit

$$M(\omega) := \lim_n M_n(\omega)$$

exists (it can be $+\infty$).
In order to identify the limit, first consider the case $x_F < \infty$. We have, for any $x < x_F$

$$P\{M_n \le x\} = \big(F(x)\big)^n \to 0,$$

so, M_n converges in law to x_F (observe that $P\{M_n \le x_F\} = 1$).
In the case $x_F = \infty$, we have that

$$P\{M = \infty\} = \lim_{K\to\infty} P\{M > K\}.$$

But,

$$P\{M > K\} \ge P\{M_n > K\} = 1 - F(K)^n \longrightarrow 1, \text{ as } n \to \infty.$$

The above result suggests us that in order to obtain a non degenerate limit, we must suitably normalize the random variables M_n.
So, we will study the possible limit laws of

$$\frac{M_n - b_n}{a_n},$$

that is, the limit of

$$\left(F(a_n x + b_n)\right)^n$$

for some convenient sequences $\{a_n\}$ and $\{b_n\}$, that are called norming (or normalizing) constants.

The main result of EVT: Fisher-Tippet theorem

Theorem

If there exist norming constants $a_n > 0$ and $b_n \in \mathbb{R}$ such that

$$\frac{M_n - b_n}{a_n} \longrightarrow H \ \ \textit{in law}$$

then H belongs to the type of one of the three following distribution functions.

Fréchet:

$$\Phi_\alpha(x) = \begin{cases} \exp\{-x^{-\alpha}\}, & x > 0 \\ 0 & x < 0. \end{cases} \qquad \alpha > 0$$

Weybull-type:

$$\Psi_\alpha(x) = \begin{cases} \exp\{-(-x^{-\alpha})\}, & x < 0 \\ 1 & x \geq 0. \end{cases} \qquad (\alpha > 0)$$

Gumbel:

$$\Lambda(x) = \exp\{-e^{-x}\}.$$

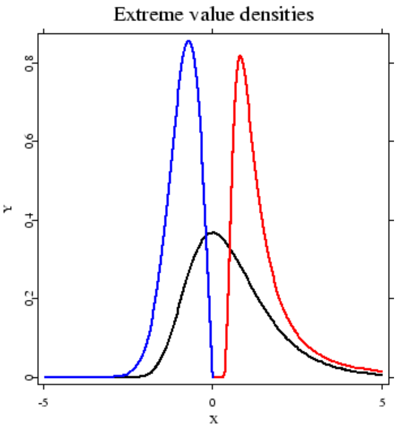

Figure: Densities of the possible limit laws for the normalized maxima. **Black line:** Gumbel density. **Red line:** Fréchet density.
Blue line: the Weibull-type density.

Sketch of the proof

- We have that for any $t > 0$

$$F^{[nt]}(a_{[nt]}x + b_{[nt]}) \to H(x).$$

However

$$F^{[nt]}(a_n x + b_n) = \left(F^n(a_n x + b_n)\right)^{[nt]/n} \to H^t(x).$$

- This implies that there exist functions $\gamma(t) > 0$, $\delta(t)$ satisfying that

$$\lim_{n\to\infty} \frac{a_n}{a_{[nt]}} = \gamma(t), \quad \lim_{n\to\infty} \frac{b_n - b_{[nt]}}{a_{[nt]}} = \delta(t)$$

and

$$H^t(x) = H\big(\gamma(t)x + \delta(t)\big).$$

- From this, it is not difficult to deduce the following functional equations for the functions γ and δ:

$$\gamma(st) = \gamma(s)\gamma(t), \quad \delta(st) = \gamma(t)\delta(s) + \delta(t).$$

The solution of these functional equations leads to the three types Λ, Φ_α and Ψ_α.

The domain of attraction of the Gumbel law

We are interested in the limit law of the maxima of standard normal distributions.
As in Central Limit Theorems, the three possible limit laws of the maxima have their corresponding maximum domain of attraction.
It is known that the normal law belongs to the maximum domain of attraction of the Gumbel distribution, this is because of the following results.

Definition

Von Mises function: Let F be a distribution function with right endpoint $x_F \leq \infty$. Suppose that there exists some $z \in [-\infty, x_F)$ such that F has the representation

$$1 - F(x) = C \exp\left\{-\int_z^x \frac{1}{D(t)} dt\right\}, \; z < x < x_F,$$

where C is some positive constant, D is a positive absolutely continuous function such that

$$\lim_{x \to x_F} D'(x) = 0.$$

In this case, we say that F is a Von Misses function.

Theorem

Suppose that the distribution function F is a von Mises function. Then, it belongs to the domain of attraction of the Gumbel law.
Moreover, as norming constants, one can take

$$b_n = F^{\leftarrow}(1 - \frac{1}{n}), \quad a_n = D(b_n).$$

Here $F^{\leftarrow}$ denotes the generalized inverse function of F.

Representation of the standard normal distribution function as a von Mises function

We will use the habitual notations

$$\phi(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}$$

$$\Phi(x) = \int_{-\infty}^{x} \phi(t) dt.$$

We have that, in this case $x_\Phi = +\infty$ and

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

this gives that our function D is

$$D(t) = \frac{1 - \Phi(t)}{\phi(t)}.$$

It is not difficult to see that

$$\lim_{x\to\infty} D'(x) = 0.$$

This function D is the so called Mills ratio of the normal standard distribution, that in what follows se will denote by M.
Applying the above theorem with this representation, we can take

$$b_n = \Phi^{-1}(1 - \frac{1}{n}) \qquad a_n = M(b_n) = \frac{1 - \Phi(b_n)}{\phi(b_n)}.$$

In order to have other possibilities for the choice of the norming constants, we can use the following general result on convergence in law, adapted to our situation.

Proposition

Suppose that $\dfrac{M_n - b_n}{a_n} \to G$ in distribution. If the sequences $\{a'_n,\, n \geq 1\}$ and $\{b'_n,\, n \geq 1\}$ satisfy

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

then

$$\frac{M_n - b'_n}{a'_n} \to G \quad \text{in distribution.}$$

(Here, we are denoting by G a random variable with Gumbel's law).

And we can also make use of this other property.

Proposition

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.

In order to have other possibilities for the choice of the norming constants, we can use the following general result on convergence in law, adapted to our situation.

Proposition

Suppose that $\dfrac{M_n - b_n}{a_n} \to G$ in distribution. If the sequences $\{a'_n, \, n \geq 1\}$ and $\{b'_n, \, n \geq 1\}$ satisfy

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

then

$$\frac{M_n - b'_n}{a'_n} \to G \quad \text{in distribution.}$$

(Here, we are denoting by G a random variable with Gumbel's law).
And we can also make use of this other property.

Proposition

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.

Properties of the Mills Ratio of the standard normal distribution

In order to obtain other norming constants and for the proof of our result we need to study the behavior of the Mills ratio.
On one hand, it is easy to see that

$$M'(x) = xM(x) - 1,$$

and from this we obtain that for any $n \geq 1$

$$M^{(n)}(x) = P_n(x)M(x) - Q_n(x),$$

with P_n polynomial of degree n, Q_n with degree $n-1$ and having both non negative coefficients. This fact allows to find the asymptotical behavior of M (as $x \to \infty$).

Properties of the Mills Ratio

On other hand,

$$M(x) = \frac{1-\Phi(x)}{\phi(x)} = \frac{\int_x^\infty e^{-t^2/2}\,dt}{e^{-x^2/2}} = \int_0^\infty e^{-xt}\, e^{-t^2/2} dt.$$

This expression gives the sign of the derivatives of M:

$$(-1)^n\, M^{(n)}(x) > 0, \ \text{ for all } x.$$

Using this, one have

$$M'(x) = xM(x) - 1 < 0$$
$$M''(x) = (1+x^2)M(x) - x > 0$$

that imply

$$\frac{\mathbf{x}}{\mathbf{x^2+1}} < \mathbf{M(x)} < \frac{\mathbf{1}}{\mathbf{x}},$$

This gives the asymptotics

$$M(x) \sim \frac{1}{x}, \quad \text{as } x \to \infty,$$

That is,

$$1 - \Phi(x) \sim \frac{1}{x}\,\phi(x) = \frac{1}{x}\frac{1}{\sqrt{2\pi}}e^{-x^2/2} \quad \text{as } x \to \infty.$$

Other norming constants

We can take a distribution function F such that, for some $x_0 > 0$, it satisfies

$$1 - F(x) = \frac{1}{x}\frac{1}{\sqrt{2\pi}}e^{-x^2/2}, \quad x \geq x_0.$$

So, we have that

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

By applying the general results on normalizing constants stated above, we can take b_n^* such that $F(b_n^*) = 1 - \frac{1}{n}$.

Or equivalently, b_n^* satisfying

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

and

$$a_n^*=\frac{1-F(b_n^*)}{F'(b_n^*)}$$

(Observe that b_n and b_n^* tend to ∞ as $n\to\infty$). Some direct computations give that

$$\frac{1-F(x)}{F'(x)}=\frac{x}{x^2+1},\quad x\geq x_0.$$

So, we can take

$$a_n^*=\frac{b_n^*}{(b_n^*)^2+1},$$

or even

$$a_n^*=\frac{1}{b_n^*}.$$

Hall's results on the velocity of convergence

Peter Hall proved that taking b_n^* such that

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

the following result holds.

Theorem

For $n \geq 2$,

$$\frac{C'}{\log n} < \sup_{x\in\mathbb{R}}|\Phi^n(a_n^*x+b_n^*)-\Lambda(x)| < \frac{C}{\log n}, \tag{1}$$

with $C = 3$

He proved also that the rate of convergence cannot be improved by choosing a different sequence of norming constants.

Notice that if $2 \leq n \leq 20$, then $3/\log n > 1$, so the upper bound in (1) gives no information.

It should also be remarked that Hall points out that his constant C in (1) can be decreased to 0.91 when $n \geq 10^6$.

Hall's results on the velocity of convergence

Peter Hall proved that taking b_n^* such that

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

the following result holds.

Theorem

For $n \geq 2$,

$$\frac{C'}{\log n} < \sup_{x\in\mathbb{R}} |\Phi^n(a_n^* x + b_n^*) - \Lambda(x)| < \frac{C}{\log n}, \qquad (1)$$

with $C = 3$

He proved also that the rate of convergence cannot be improved by choosing a different sequence of norming constants.

Notice that if $2 \leq n \leq 20$, then $3/\log n > 1$, so the upper bound in (1) gives no information.

It should also be remarked that Hall points out that his constant C in (1) can be decreased to 0.91 when $n \geq 10^6$.

Hall's results on the velocity of convergence

Peter Hall proved that taking b_n^* such that

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

the following result holds.

Theorem

For $n \geq 2$,

$$\frac{C'}{\log n} < \sup_{x\in\mathbb{R}} |\Phi^n(a_n^* x + b_n^*) - \Lambda(x)| < \frac{C}{\log n}, \tag{1}$$

with $C = 3$

He proved also that the rate of convergence cannot be improved by choosing a different sequence of norming constants.

Notice that if $2 \leq n \leq 20$, then $3/\log n > 1$, so the upper bound in (1) gives no information.

It should also be remarked that Hall points out that his constant C in (1) can be decreased to 0.91 when $n \geq 10^6$.

Hall's results on the velocity of convergence

Peter Hall proved that taking b_n^* such that

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

the following result holds.

Theorem

For $n \geq 2$,

$$\frac{C'}{\log n} < \sup_{x\in\mathbb{R}} |\Phi^n(a_n^* x + b_n^*) - \Lambda(x)| < \frac{C}{\log n}, \qquad (1)$$

with $C = 3$

He proved also that the rate of convergence cannot be improved by choosing a different sequence of norming constants.
Notice that if $2 \leq n \leq 20$, then $3/\log n > 1$, so the upper bound in (1) gives no information.
It should also be remarked that Hall points out that his constant C in (1) can be decreased to 0.91 when $n \geq 10^6$.

Main result

Our choice of norming constants is:

$$b_n = \Phi^{-1}(1 - \frac{1}{n}), \quad a_n = \frac{b_n}{b_n^2 + 1}.$$

Theorem

Given $n_0 \geq 5$, for all $n \geq n_0$ it holds that

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

with

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

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.

We can give some bounds for $\{b_n^2\}$ (see the next proposition) 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: Several upper approximations for $C(n_0)$ and $\widetilde{C}(n_0)$.

Motivation

With

$$Y_n = \frac{M_n - b_n}{a_n} \qquad \text{with our constants}$$

$$Y_n^* = \frac{M_n - b_n^*}{a_n^*} \qquad \text{with Hall's constants}$$

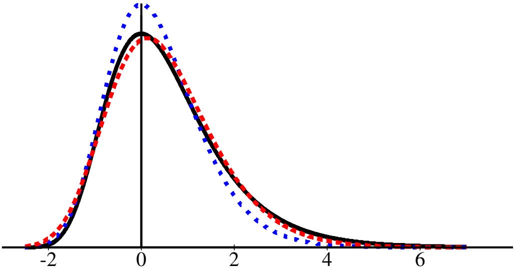

Figure: Gumbel density and density of the maximum of 100 standard Gaussian random variables with different norming constants. Solid line: Gumbel density. Dotted blue line: Density of Y_n^*. Dashed red line: Density of Y_n.

Some ideas of the proof of the main result

Proposition

Let $b_n = \Phi^{-1}(1 - n^{-1})$. For each $n \geq 2$ the following inequalities hold:

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

Proof: Recall that

$$b_n = \Phi^{-1}(1 - \tfrac{1}{n}).$$

We will only see the right hand side inequality in (2). First of all, observe that for $n = 2$ we have that $b_2 = 0$, while $2 \log 2 > 0$. So, we consider the case $n \geq 3$.
We will prove that for $n \geq 3$,

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

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 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

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

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

Computing the difference $\Phi^n(a_n + b_n x) - \Lambda(x)$

Recall the expression of the Mills ratio

$$M(t) = \frac{1 - \Phi(t)}{\phi(t)}$$

and denote

$$V(t) = \frac{1}{M(t)} = \frac{\phi(t)}{1 - \Phi(t)}.$$

Recall also the expression of Φ as a Von Mises function

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

Using this, we have

$$1 - \Phi(a_n x + b_n) = \exp\left\{-\int_{-\infty}^{b_n} V(t)\,dt\right\} \exp\left\{-\int_{b_n}^{a_n x + b_n} V(t)\,dt\right\}$$

$$= \frac{1}{n} \exp\left\{-\int_{b_n}^{a_n x + b_n} V(t)\,dt\right\},$$

where we have used that $\Phi(b_n) = 1 - 1/n$.

Computing the difference $\Phi^n(a_n + b_n x) - \Lambda(x)$

Recall the expression of the Mills ratio

$$M(t) = \frac{1 - \Phi(t)}{\phi(t)}$$

and denote

$$V(t) = \frac{1}{M(t)} = \frac{\phi(t)}{1 - \Phi(t)}.$$

Recall also the expression of Φ as a Von Mises function

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

Using this, we have

$$1 - \Phi(a_n x + b_n) = \exp\left\{-\int_{-\infty}^{b_n} V(t)\,dt\right\} \exp\left\{-\int_{b_n}^{a_n x + b_n} V(t)\,dt\right\}$$

$$= \frac{1}{n} \exp\left\{-\int_{b_n}^{a_n x + b_n} V(t)\,dt\right\},$$

where we have used that $\Phi(b_n) = 1 - 1/n$.

Computing the difference $\Phi^n(a_n + b_n x) - \Lambda(x)$

Recall the expression of the Mills ratio

$$M(t) = \frac{1 - \Phi(t)}{\phi(t)}$$

and denote

$$V(t) = \frac{1}{M(t)} = \frac{\phi(t)}{1 - \Phi(t)}.$$

Recall also the expression of Φ as a Von Mises function

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

Using this, we have

$$1 - \Phi(a_n x + b_n) = \exp\left\{-\int_{-\infty}^{b_n} V(t)\, dt\right\} \exp\left\{-\int_{b_n}^{a_n x + b_n} V(t)\, dt\right\}$$

$$= \frac{1}{n} \exp\left\{-\int_{b_n}^{a_n x + b_n} V(t)\, dt\right\},$$

where we have used that $\Phi(b_n) = 1 - 1/n$.

Computing the difference $\Phi^n(a_n + b_n x) - \Lambda(x)$

Then,

$$\Phi(a_n x + b_n) = 1 - \frac{1}{n} \exp\left\{ - \int_{b_n}^{a_n x + b_n} V(t)\, dt \right\}, \tag{3}$$

and so

$$\log \Phi^n(a_n' x + b_n) = n \log\left(1 - \frac{1}{n} e^{-I_n(x)}\right) = -e^{-I_n(x)} - nS_n(x), \tag{4}$$

where we are denoting by

$$I_n(x) = \int_{b_n}^{a_n x + b_n} V(t)\, dt,$$

$S_n(x)$ is the remaining term of first order Taylor's development of

$$\log(1-u),$$

for $u \in (0,1)$, with

$$u = \frac{1}{n} e^{-I_n(x)}$$

(Observe that, by (3), $e^{-I_n(x)}/n \in (0,1)$).

Computing the difference $\Phi^n(a_n + b_n x) - \Lambda(x)$

Hence,

$$\begin{aligned}\Phi^n(a_n x + b_n) - \Lambda(x) &= \exp\log\Phi^n(a_n x + b_n) - \Lambda(x)\\ &= \exp\Big(-e^{-I_n(x)} - nS_n(x)\Big) - \Lambda(x)\\ &= e^{-nS_n(x)}\Lambda\big(I_n(x)\big) - \Lambda(x).\end{aligned}$$

Adding and subtracting the term $e^{-nS_n(x)}\Lambda(x)$ we arrive at:

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

Recall that

$$I_n(x) = \int_{b_n}^{a_n x + b_n} V(t)\,dt.$$

On the other hand, it can be seen that

$$0 < S_n(x) < \frac{C_n(x)^2}{2(1 - C_n(x))}, \quad \text{with } C_n(x) = \frac{1}{n}e^{-I_n(x)}.$$

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

$$I_n(x) = \int_{b_n}^{a_n x + b_n} V(t)\,dt; \qquad 0 < S_n(x) < \frac{C_n(x)^2}{2(1 - C_n(x))}, \qquad C_n(x) = \frac{1}{n} e^{-I_n(x)}.$$

We divide the proof of our main Theorem in two cases: $x \geq 0$ and $x < 0$.
The case $x \geq 0$ is easier, since in this case

$$I_n(x) \geq 0,$$

that implies

$$0 < C_n(x) \leq 1/n$$

and then

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

Recall

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

$$I_n(x) = \int_{b_n}^{a_n x + b_n} V(t)\,dt; \qquad 0 < S_n(x) < \frac{C_n(x)^2}{2(1 - C_n(x))}, \qquad C_n(x) = \frac{1}{n} e^{-I_n(x)}.$$

So, the more difficult term to study is

$$|\Lambda(I_n(x)) - \Lambda(x)|.$$

For this, we use that

$$\frac{x}{x^2+1} < M(x) < \frac{1}{x},$$

and consequently we have

$$x < V(x) < 1 + \frac{1}{x}$$

It is necessary to separate the two cases $I_n(x) < x$ and $I_n(x) \geq x$.
For instance, if $I_n(x) < x$:

$$\begin{aligned} \left|\Lambda\big(I_n(x)\big) - \Lambda(x)\right| &= \Lambda(x) - \Lambda\big(I_n(x)\big) \leq \Lambda'\big(I_n(x)\big)\big(x - I_n(x)\big) \\ &= \Lambda\big(I_n(x)\big)e^{-I_n(x)}\big(x - I_n(x)\big) \leq e^{x-I_n(x)}e^{-x}\big(x - I_n(x)\big). \end{aligned} \tag{5}$$

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 $V(t) > t$.

It is necessary to separate the two cases $I_n(x) < x$ and $I_n(x) \geq x$.
For instance, if $I_n(x) < x$:

$$\Big|\Lambda\big(I_n(x)\big) - \Lambda(x)\Big| = \Lambda(x) - \Lambda\big(I_n(x)\big) \leq \Lambda'\big(I_n(x)\big)\big(x - I_n(x)\big)$$
$$= \Lambda\big(I_n(x)\big)e^{-I_n(x)}\big(x - I_n(x)\big) \leq e^{x-I_n(x)}e^{-x}\big(x - I_n(x)\big). \qquad (5)$$

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 $V(t) > t$.

It is necessary to separate the two cases $I_n(x) < x$ and $I_n(x) \geq x$.
For instance, if $I_n(x) < x$:

$$\Big|\Lambda\big(I_n(x)\big) - \Lambda(x)\Big| = \Lambda(x) - \Lambda\big(I_n(x)\big) \leq \Lambda'\big(I_n(x)\big)\big(x - I_n(x)\big)$$
$$= \Lambda\big(I_n(x)\big)e^{-I_n(x)}\big(x - I_n(x)\big) \leq e^{x-I_n(x)}e^{-x}\big(x - I_n(x)\big). \qquad (5)$$

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 $V(t) > t$.

Hence

$$\begin{aligned}\left|\Lambda\left(l_n(x)\right)-\Lambda(x)\right| &\le e^{x-l_n(x)}e^{-x}\left(x-l_n(x)\right)\\ &\le e^{\frac{x}{b_n^2+1}}e^{-x}\frac{x}{b_n^2+1}=e^{-\frac{b_n^2x}{b_n^2+1}}\frac{x}{b_n^2+1}\\ &=e^{-\frac{b_n^2x}{b_n^2+1}}\frac{b_n^2x}{b_n^2+1}\frac{1}{b_n^2}\le \max_{y\in[0,\infty)}\left\{e^{-y}y\right\}\frac{1}{b_n^2}=\frac{1}{eb_n^2}.\end{aligned}$$

Some words about the case $x < 0$

We have divided the values of x according whether

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

In the case $x \in (-\infty, -b_n/a_n)$, we have that

$$x < 0, \quad \text{and} \quad a_n x + b_n < 0$$

and this allows to bound separately

$$\Lambda(x) \quad \text{and} \quad \Phi^n(a_n x + b_n).$$

And this is not difficult.

In the case $x \in [-b_n/a_n, -1.25 \log b_n]$ we have

$$x < 0, \quad a_n x + b_n \geq 0.$$

Nevertheless, we can also study separately

$$\Lambda(x) \quad \text{and} \quad \Phi^n(a_n x + b_n).$$

HARD!

Case $x \in (-1.25 \log n, 0)$.
We now must use once more

$$\Phi^n(a_n x + b_n) - \Lambda(x) = e^{-nS_n(x)}\Big(\Lambda\big(l_n(x)\big) - \Lambda(x)\Big) + \Lambda(x)\left(e^{-nS_n(x)} - 1\right)$$

$$l_n(x) = \int_{b_n}^{a_n x + b_n} V(t)\,dt; \qquad 0 < S_n(x) < \frac{C_n(x)^2}{2(1 - C_n(x))}, \qquad C_n(x) = \frac{1}{n} e^{-l_n(x)}.$$

We comment only the study of first summand of the expression of the difference $\Phi^n(a_n x + b_n) - \Lambda(x)$. The problem, here is that $-x$ and $-l_n(x)$ are now positive. But we can see that

$$-l_n(x) \leq -x.$$

Case $x \in (-1.25 \log n, 0)$.
We now must use once more

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

$$I_n(x) = \int_{b_n}^{a_n x + b_n} V(t)\,dt; \qquad 0 < S_n(x) < \frac{C_n(x)^2}{2(1 - C_n(x))}, \qquad C_n(x) = \frac{1}{n} e^{-I_n(x)}.$$

We comment only the study of first summand of the expression of the difference $\Phi^n(a_n x + b_n) - \Lambda(x)$. The problem, here is that $-x$ and $-I_n(x)$ are now positive. But we can see that

$$-I_n(x) \le -x.$$

Case $x \in (-1.25 \log n, 0)$.
We now must use once more

$$\Phi^n(a_n x + b_n) - \Lambda(x) = e^{-nS_n(x)} \Big(\Lambda\big(l_n(x)\big) - \Lambda(x) \Big) + \Lambda(x) \left(e^{-nS_n(x)} - 1 \right)$$

$$l_n(x) = \int_{b_n}^{a_n x + b_n} V(t)\, dt; \qquad 0 < S_n(x) < \frac{C_n(x)^2}{2(1 - C_n(x))}, \quad C_n(x) = \frac{1}{n} e^{-l_n(x)}.$$

We comment only the study of first summand of the expression of the difference $\Phi^n(a_n x + b_n) - \Lambda(x)$. The problem, here is that $-x$ and $-l_n(x)$ are now positive.
But we can see that

$$-l_n(x) \le -x.$$

Case $x \in (-1.25 \log n, 0)$.
We now must use once more

$$\Phi^n(a_n x + b_n) - \Lambda(x) = e^{-nS_n(x)}\Big(\Lambda\big(l_n(x)\big) - \Lambda(x)\Big) + \Lambda(x)\left(e^{-nS_n(x)} - 1\right)$$

$$l_n(x) = \int_{b_n}^{a_n x + b_n} V(t)\,dt; \qquad 0 < S_n(x) < \frac{C_n(x)^2}{2(1 - C_n(x))}, \qquad C_n(x) = \frac{1}{n} e^{-l_n(x)}.$$

We comment only the study of first summand of the expression of the difference $\Phi^n(a_n x + b_n) - \Lambda(x)$. The problem, here is that $-x$ and $-l_n(x)$ are now positive. But we can see that

$$-l_n(x) \leq -x.$$

Using also that $-x \leq 1.25 \log b_n$, and then $b_n^2 \geq \exp(-8x/5)$, and after some not difficult computations, one can see that

$$\Lambda\big(l_n(x)\big) - \Lambda(x) \leq \frac{1}{b_n^2}\Big(-x + \frac{x^2}{2}\Big) \exp\Big\{-x - e^{-x} + \Big(-x + \frac{x^2}{2}\Big)e^{3x/5}\Big\}$$
$$= \frac{1}{b_n^2}\, Q(x)$$

We can prove that

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

Idea about how to prove this:

- Draw this function with some software and see that there is a unique global maximum x^* and that $Q(x^*)$ seems to be less than 0.63.
- Prove that $Q'(x)$ has an only zero that corresponds to a maximum. (This is the most difficult part!)
- In our graphic of Q we see that x^* is around -1.05. Then, using Bolzano's theorem we see that $x^* \in (\underline{x}, \overline{x}) = (-1.051, -1.050)$
- Finally:

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

We can prove that

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

Idea about how to prove this:

- Draw this function with some software and see that there is a unique global maximum x^* and that $Q(x^*)$ seems to be less than 0.63.
- Prove that $Q'(x)$ has an only zero that corresponds to a maximum. (This is the most difficult part!)
- In our graphic of Q we see that x^* is around -1.05. Then, using Bolzano's theorem we see that $x^* \in (\underline{x}, \overline{x}) = (-1.051, -1.050)$
- Finally:

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

We can prove that

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

Idea about how to prove this:

- Draw this function with some software and see that there is a unique global maximum x^* and that $Q(x^*)$ seems to be less than 0.63.
- Prove that $Q'(x)$ has an only zero that corresponds to a maximum. (This is the most difficult part!)
- In our graphic of Q we see that x^* is around -1.05. Then, using Bolzano's theorem we see that $x^* \in (\underline{x}, \overline{x}) = (-1.051, -1.050)$
- Finally:

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

We can prove that

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

Idea about how to prove this:

- Draw this function with some software and see that there is a unique global maximum x^* and that $Q(x^*)$ seems to be less than 0.63.
- Prove that $Q'(x)$ has an only zero that corresponds to a maximum. (This is the most difficult part!)
- In our graphic of Q we see that x^* is around -1.05. Then, using Bolzano's theorem we see that $x^* \in (\underline{x}, \overline{x}) = (-1.051, -1.050)$
- Finally:

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

Explicit expressions for the norming constants

Recall that

$$b_n = \Phi^{-1}(1 - \frac{1}{n}), \qquad a_n = \frac{b_n}{b_n^2 + 1}.$$

As n is very big, we must find the inverse of the distribution function of the standard normal distribution at points near to 1. This is a numerical problem.
This fact leads to the necessity of finding some useful asymptotic expansions of the b_n's.

Proposition

It holds that

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

where

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

Explicit expressions for the norming constants

Its proof uses sharper bounds of the Mills ratio

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

with

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

that implies

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

We use also the asymptotic development of Lambert-type functions. That is, the asymptotic development, as $t \to \infty$, of functions $g(t)$ that are solutions of the equation

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

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.

Explicit expressions for the norming constants

Finally, we propose the following approximations of b_n that work better for moderate n's.

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

(We change the term $\log\log\left(n^2/(2\pi)\right)$ of the numerator of the last fraction in the expression of $\overline{\beta}_n$ by $\log\left(\log(n^2)+1/2\right)$.)

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

Table: Comparison of the proposed constants $\overline{\beta}_n$ and β_n with b_n.

Explicit expressions for the norming constants

Finally, we propose the following approximations of b_n that work better for moderate n's.

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

(We change the term $\log\log\left(n^2/(2\pi)\right)$ of the numerator of the last fraction in the expression of $\overline{\beta}_n$ by $\log\left(\log(n^2)+1/2\right)$.)

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

Table: Comparison of the proposed constants $\overline{\beta}_n$ and β_n with b_n.

Bibliography

- DE BRUIJN, N.G., *Asymptotic Methods in Analysis*, Dover, New York, 1981.
- 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.
- GASULL, A. AND UTZET, F., Approximating Mills ratio. *J. Math. Anal. Appl.* **420** 1832–1853.
- GNEDENKO, B.V., Sur la distribution limite du terme maximum d une série aléatoire. *Ann. Math.* **44** (1943) 423–453.
- HALL, P., On the rate of convergence of normal extremes. *J. Appl. Probab.* **16** (1979) 433–439.
- RESNICK, S.I., *Extreme Values, Regular Variation, and Point Processes*, Springer, Berlin, 1987.