ON THE RATE OF CONVERGENCE OF NORMAL EXTREMES

PETER HALL,* University of Melbourne

Abstract

Let $Y_n$ denote the largest of $n$ independent $N(0,1)$ variables. It is shown that if the constants $a_n$ and $b_n$ are chosen in an optimal way then the rate of convergence of $(Y_n - b_n)/a_n$ to the extreme value distribution $\exp(-e^{-x})$, as measured by the supremum metric or the Lévy metric, is proportional to $1/\log n$.

NORMAL DISTRIBUTION; EXTREME VALUE; EXTREME VALUE DISTRIBUTION; RATE OF CONVERGENCE

Let $X_1, X_2, \ldots$ be independent $N(0,1)$ variables and let $Y_n = \max_{i \leq n} X_i$ denote the largest of the first $n$. For suitable constants $a_n$ and $b_n$, $(Y_n - b_n)/a_n$ has the limiting distribution $\Lambda$ defined by $\Lambda(x) = \exp(-e^{-x})$. That is to say, as $n \to \infty$

$$\Phi^{\circ}(a_n x + b_n) \to \Lambda(x), \quad -\infty < x < \infty,$$

where $\Phi$ is the distribution function of a standard normal variable.

The limiting behaviour of extreme values was first elucidated by Fisher and Tippett (1928), and they remarked that in the case of normal extremes the limit is approached extremely slowly. In this respect the normal distribution seems to be peculiar among the more common distributions in the domain of attraction of $\Lambda$. For example, it is easily seen that the largest of $n$ independent negative exponential variables converges to the extreme value distribution at a rate $1/n$. Our aim in this note is to show that the rate of convergence in the case of normal extremes is $1/\log n$.

The most natural way of defining the norming constants $a_n$ and $b_n$ is to let $b_n$ be the solution of the equation

$$(1a) \quad 2\pi b_n^2 \exp(b_n^2) = n^2$$

and set

$$(1b) \quad a_n = b_n^{-1}$$

Received 31 January 1978; revision received 23 May 1978.

* Present address: Department of Statistics, S.G.S., The Australian National University, P.O. Box 4, Canberra, A.C.T. 2600, Australia.
If $\alpha_n$ and $\beta_n$ are any other suitable constants then

$$\frac{\alpha_n}{a_n} \to 1 \quad \text{and} \quad \frac{(\beta_n - b_n)}{a_n} \to 0$$

(Gnedenko (1943)). The solution of (1a) may be expressed as

$$b_n = (2\log n)^{1/2} - \frac{1}{2}(\log \log n + \log 4\pi)/(2\log n)^{1/2} + O(1/\log n)$$

(Cramér (1946), p. 374), and it is easy to see that for $n \geq 2$,

$$2\log n - (\log \log n + \log 4\pi) < b_n^2 < 2\log n.$$

If $\alpha_n$ and $\beta_n$ are defined by

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

and if $a_n$ and $b_n$ are defined as in (1), then (2) implies that $\alpha_n/a_n \to 1$ and $(\beta_n - b_n)/a_n \to 0$, so that $\alpha_n$ and $\beta_n$ are suitable norming constants. (In this regard we should point out an error in David (1970), p. 209, where it is incorrectly stated that suitable norming constants are $\beta_n = \alpha_n^{-1} = (2\log n)^{1/2}$.)

**Theorem.** There exist positive constants $C_1$ and $C_2$, independent of $n$, such that for all $n$,

$$\frac{C_1}{\log n} < \sup_{-\infty < x < \infty} \left| \Phi^*(a_n x + b_n) - \Lambda(x) \right| < C_2/\log n,$$

where $a_n$ and $b_n$ are defined by (1). The constant $C_2$ may be taken equal to 3. The rate of convergence cannot be improved by choosing a different sequence of norming constants. In fact, with the norming constants defined in (4) the rate of convergence is not better than $(\log \log n)^2/\log n$.

**Remarks.** The supremum metric used in (5) may be replaced by the Lévy metric. Let $\rho_n$ and $\lambda_n$ denote, respectively, the uniform and Lévy distances between the distribution functions $\Phi^*(a_n \cdot + b_n)$ and $\Lambda$. Then

$$\lambda_n \leq \rho_n \leq (\lambda_n + e^{-1})\lambda_n$$

(Zolotarev (1967), Lemma 2), and so $\rho_n$ may be replaced by $\lambda_n$ in the inequalities (5).

If the uniform bound in (5) is only required for large values of $n$ then the constant $C_2$ may be assigned a value considerably less than 3. For example, if we are considering $n \geq 10^6$ then since $b_{10^6} = 4.7615$, the bounds in (15)–(19) may be replaced by 1.11, 0.14, 0.43, $3.4 \times 10^{-6}$ and $1.3 \times 10^{-3}$, respectively. Our proof now shows that $C_2 = 0.91$ will suffice.

**Proof.** For any $x > 0$ we can write

$$1 - \Phi(x) = x^{-1}(2\pi)^{-1/2}e^{-x^2/2} - r(x)$$

$$= x^{-1}(2\pi)^{-1/2}e^{-x^2/2}(1 - x^{-2}) + s(x)$$

This content downloaded from 133.23.3.140 on Tue, 28 Jan 2020 09:30:21 UTC
All use subject to https://about.jstor.org/terms
where
\[ 0 < r(x) = \int_x^\infty t^{-4}(2\pi)^{-\frac{1}{2}}e^{-\frac{1}{4}t^2}dt < x^{-3}(2\pi)^{-\frac{1}{2}}e^{-\frac{1}{4}x^2} \]
and
\[ 0 < s(x) = \int_x^\infty t^{-4}(2\pi)^{-\frac{1}{2}}e^{-\frac{1}{4}t^2}dt < 3x^{-3}(2\pi)^{-\frac{1}{2}}e^{-\frac{1}{4}x^2} \]
(Abramowitz and Stegun (1964), p. 932). Suppose that \( \alpha_n \) and \( \beta_n \) are suitable norming constants. Then
\[ \alpha_n = a_nr_n \quad \text{and} \quad \beta_n = \delta_n a_n + b_n \]
where \( r_n \to 1 \) and \( \delta_n \to 0 \). If \( n \) is sufficiently large then \( \alpha_n x + \beta_n > 0 \), and so
\[
(\alpha_n x + \beta_n)^{-\frac{1}{2}}(2\pi)^{-\frac{1}{4}}\exp[-\frac{1}{2}(\alpha_n x + \beta_n)^2]
\]
\[ = (2\pi b_n^2 \exp(b_n^2))^{-\frac{1}{4}}[1 + a_n^2(r_n x + \delta_n)]^{-\frac{1}{4}}\exp[-\frac{1}{2}a_n^2(r_n x + \delta_n)^2 - (r_n x + \delta_n)]
\]
\[ = n^{-1}e^{-\frac{1}{2}}[1 - a_n^2(r_n x + \delta_n) + O(a_n^4)]
\]
\[ \times [1 - \frac{1}{2}a_n^2(r_n x + \delta_n)^2 - (r_n - 1)x - \delta_n + O(a_n^4 + (r_n - 1)^2 + \delta_n^2)]
\]
\[ = n^{-1}e^{-\frac{1}{2}}[1 - a_n^2x(1 + \frac{1}{2}x) - (r_n - 1)x - \delta_n + O(a_n^4 + (r_n - 1)^2 + \delta_n^2)].
\]
Similarly,
\[ (\alpha_n x + \beta_n)^{-\frac{1}{2}}(2\pi)^{-\frac{1}{4}}\exp[-\frac{1}{2}(\alpha_n x + \beta_n)^2] = O(n^{-1}a_n^4) \]
and
\[ 1 - (\alpha_n x + \beta_n)^{-2} = 1 - a_n^2 + O(a_n^4). \]

Using these results together with (7) and (9) we deduce that
\[
\Phi^*(\alpha_n x + \beta_n) - \Lambda(x)
\]
\[ = \{1 - n^{-1}e^{-\frac{1}{2}}[1 - a_n^2(1 + x + \frac{1}{2}x^2) - (r_n - 1)x - \delta_n
\]
\[ + O(a_n^4 + (r_n - 1)^2 + \delta_n^2)]^n - \Lambda(x)
\]
\[ = \Lambda(x)e^{-\frac{1}{2}}[a_n^2(1 + x + \frac{1}{2}x^2) + (r_n - 1)x + \delta_n + O(a_n^4 + (r_n - 1)^2 + \delta_n^2)].
\]
The inequalities (3) imply that \( a_n^2 \sim 1/\log n \) and so the rate of convergence cannot be better than \( 1/\log n \). Setting \( r_n = 1 \) and \( \delta_n = 0 \) we obtain the left-hand inequality in (5).

Suppose now that the norming constants are defined as in (4). By making two successive applications of the Newton–Rhapson approximation method we can extend the series expansion in (2):
\[
b_n = (2\log n)^\frac{1}{2} - \frac{1}{2}(\log \log n + \log 4\pi)/(2\log n)^\frac{1}{2}
\]
\[ - [(\log \log n + \log 4\pi)^2 - 4(\log \log n + \log 4\pi)]/8(2\log n)^\frac{3}{2}
\]
\[ + O((\log \log n)^3/(\log n)^{3/2}).
\]
Therefore
\[ \delta_n = b_n (\beta_n - b_n) = (\log \log n)^2 / 16 \log n + \text{smaller order terms}. \]

It now follows from (10) that the rate of convergence is not better than \((\log \log n)^2 / \log n\).

It remains to prove that
\[ \sup_{-\infty < x \leq c} |\Phi^n(a_n x + b_n) - \Lambda(x)| < 3 / \log n. \]

Since \(3 / \log 20 > 1\) then it suffices to establish the inequality for \(n \geq 21\). The inequalities (3) imply that for \(n \geq 21\), \(b_n^2 > 0.8 \log n\), and so it suffices to prove that for \(n \geq 21\),
\[ \sup_{-\infty < x < c_n} |\Phi^n(a_n x + b_n) - \Lambda(x)| < 2.4 a_n^2. \]

We shall do this in three parts, establishing that
\[ \sup_{0 \leq x \leq c_n} |\Phi^n(a_n x + b_n) - \Lambda(x)| < 1.57 a_n^2, \]

\[ \sup_{-c_n < x < 0} |\Phi^n(a_n x + b_n) - \Lambda(x)| < 1.02 a_n^2, \]

and
\[ \sup_{-\infty < x \leq -c_n} |\Phi^n(a_n x + b_n) - \Lambda(x)| < 2.08 a_n^2, \]

where \(c_n = \log \log b_n^2 (> 0 \text{ for } n \geq 21)\). The following bounds are easily obtained:
\[ 1.76 < b_n < 1.77, \]

\[ \sup_{n \geq 21} (1 - a_n^2 c_n)^{-1} < 1.11, \]

\[ \sup_{n \geq 21} a_n^2 \log b_n < 0.37, \]

\[ \sup_{n \geq 21} a_n^2 (\log b_n^2)^2 < 0.55, \]

\[ \sup_{n \geq 21} n^{-1} \log b_n < 0.087 \]

and
\[ \sup_{n \geq 21} b_n \exp(- \frac{1}{2} b_n^2) < 1.16. \]

((14) follows from (1a), (18) follows from (3), and (15), (16), (17) and (19) are...
obtained by bounding the functions \(x^{-1}\log \log x, x^{-1}\log x, x^{-1}(\log x)^2\) and \(x \cdot e^{-\frac{1}{2}x^2}\), respectively.)

Suppose first that \(x > -c_n\) and let \(\Psi_n(x) = 1 - \Phi(a_nx + b_n)\). Then

\[
n \log \Phi(a_nx + b_n) = n \log[1 - \Psi_n(x)] = -n\Psi_n(x) - R_n(x)
\]

where

\[
0 < R_n(x) \leq n\Psi_n(x)/2[1 - \Psi_n(x)].
\]

Proceeding as before we deduce from (6) and (8) that

\[
\Psi_n(x) < \Psi_n(-c_n) < n^{-1}(1 - a_n^2c_n)^{-1}\exp(c_n - \frac{1}{2}a_n^2c_n^2)
\]

\[
< n^{-1}(1 - a_n^2c_n)^{-1}\log b_n < 0.097.
\]

(Use (15) and (18).) From (20) and the definition of \(b_n\) we see that

\[
R_n(x) < n^{-1}(1 - a_n^2c_n)^{-2}b_n^2(\log b_n)^2/2(1 - 0.097)b_n^2 = (2\pi)^{-\frac{1}{2}}[a_n^2(\log b_n)^2][b_n^2\exp(-\frac{1}{2}b_n^2)(1 - a_n^2c_n)^{-2}/1.806 b_n^2
\]

\[
< 0.18a_n^2.
\]

(Use (15), (17) and (19).) It follows that for \(n \geq 21,

\[
|\exp(-R_n(x)) - 1| < 0.18a_n^2.
\]

Let \(A_n(x) = \exp[-n\Psi_n(x) + e^{-x}]\) and \(B_n(x) = \exp[-R_n(x)]\). The inequality (21) implies that

\[
|\Phi^*(a_nx + b_n) - \Lambda(x)| = \Lambda(x)|A_n(x)B_n(x) - 1|
\]

\[
\leq \Lambda(x)|B_n(x)||A_n(x) - 1| + |B_n(x) - 1|
\]

\[
< \Lambda(x)|A_n(x) - 1| + 0.18a_n^2
\]

if \(x \geq -c_n\).

(I) \(0 \leq x < \infty\). Since \(A_n(x) \to 1\) as \(x \to \infty\) and

\[
A_n'(x) = -A_n(x)e^{-x}[1 - \exp(-\frac{1}{2}a_n^2x^2)] < 0
\]

for \(x > 0\) then from (6) and (8),

\[
\sup_{x \geq 0}|A_n(x) - 1| = |A_n(0) - 1|
\]

\[
= |\exp(nr(b_n)) - 1|
\]

\[
< a_n^2\exp(a_n^2)
\]

\[
< 1.39a_n^2.
\]

Combined with (22) this establishes (11).
(II) \(-c_n < x < 0\). From (6) and (8) it follows that

\[-n\Psi_n(x) + e^{-x} = (1 + a_n^2x)^{-1}e^{-x}[ -\exp(-\frac{1}{2}a_n^2x^2)(1 - a_n^2d_n(1 + a_n^2x)^{-2}) + 1 + a_n^2x] = (1 + a_n^2x)^{-1}e^{-x}C_n(x),\]
say, where \(0 < d_n(x) < 1\). Now,

\[C_n(x) = (1 - \frac{1}{2}a_n^2x^2)(1 - a_n^2(1 - 2a_n^2x)) + 1 + a_n^2x = a_n^2[1 + x(1 - 2a_n^2) + \frac{1}{2}x^2(1 - a_n^2) + a_n^2x^3] < a_n^2(1 + \frac{1}{2}x^2)\]

for \(-c_n < x < 0\), and

\[C_n(x) > -1 + 1 + a_n^2x = -a_n^2|x|.\]

Since \(1 + \frac{1}{2}x^2 > |x|\) then

\[| -n\Psi_n(x) + e^{-x}| < a_n^2(1 - a_n^2c_n)^{-1}(1 + \frac{1}{2}x^2)e^{-x} < 1.11a_n^2(1 + \frac{1}{2}x^2)e^{-x}.\]

For \(-c_n < x < 0\),

\[a_n^2(1 + \frac{1}{2}x^2)e^{-x} < a_n^2\log b_n^2 + \frac{1}{2}a_n^2(\log b_n^2)^2 < a_n^2\log b_n^2 + \frac{1}{2}a_n^2(\log b_n^2)^2 < 0.37 + \frac{1}{2}(0.55) = 0.645.\]

(Use (16) and (17). The second inequality is obtained by observing that \(\log t > (\log \log t)^2\) for \(t > 1\).) Hence

\[\Lambda(x)|A_n(x) - 1| < 1.11a_n^2(1 + \frac{1}{2}x^2)\exp(-e^{-x} - x + 0.645 \times 1.11) < 1.11a_n^2(1 + \frac{1}{2}x^2)\exp(-1 + x - \frac{1}{2}x^2 - x + 0.645 \times 1.11) < 0.84a_n^2(1 + \frac{1}{2}x^2)e^{-|x^2|} < 0.84a_n^2\]

since the function \((1 + \frac{1}{2}t^2)e^{-|t^2|}\) is dominated by 1. Combined with (22) this gives (12).

(III) \(-\infty < x < -c_n\). In this case

\[\Lambda(x) \leq \Lambda(-c_n) = a_n^2\]

and from Equations (7) and (9) we deduce that
On the rate of convergence of normal extremes

\[ \Phi^n(a_nx + b_n) \leq \Phi^n(b_n - a_n c_n) \]
\[ \leq [1 - (b_n - a_n c_n)^{-1}(1 - a_n c_n)^{-2}(2\pi)^{-\frac{1}{2}}\exp(-\frac{1}{2}(b_n - a_n c_n)^2)]^n \]
\[ = [1 - n^{-1}(1 - a_n^2 c_n)^{-1}(1 - a_n^2 c_n)^{-2} \exp(c_n - \frac{1}{2} a_n^2 c_n^2)]^n \]
\[ < \exp[-(1 - a_n^2 c_n)^{-1}(1 - a_n^2(1 - a_n^2 c_n)^{-2}) \exp(c_n - \frac{1}{2} a_n^2 c_n^2)]. \]

Now,
\[ (1 - a_n^2 c_n)^{-1}(1 - a_n^2(1 - a_n^2 c_n)^{-2}) \exp(-\frac{1}{2} a_n^2 c_n^2) \]
\[ > (1 - a_n^2(1 - a_n^2 c_n)^{-2})(1 - \frac{1}{2} a_n^2 c_n^2) \]
\[ > 1 - a_n^2 \frac{1}{2} c_n^2 + (1 - a_n^2 c_n)^{-2}. \]

Hence for \( n \geq 21, \)
\[ \Phi^n(a_nx + b_n) < a_n^2 \exp\left[a_n^2 (\log b_n^2)(\frac{1}{2} c_n^2 + (1 - a_n^2 c_n)^{-2})\right] \]
\[ < a_n^2 \exp\left[\frac{1}{2}(0.55) + (0.37)(1.11)^2\right] < 2.08 a_n^2. \]

(Use (15)–(17).) Combined with (23) this implies (13) and completes the proof.

Acknowledgement

I am grateful to Mr G. Hepworth for stimulating my interest in this problem.

References


