#11

The rate of convergence in law of the maximum of an exponential sample*

by W. J. HALL** and JON A. WELLNER**

Summary We derive a uniform rate of convergence of $(1-n^{-1}x)^n$ to e^{-x} $(x \geq 0)$. It provides a uniform rate of convergence for the distribution of the largest order statistic in a sample from an exponential distribution to the "double exponential" extreme value distribution. It likewise provides a rate of convergence for the distribution of the smallest order statistic from a uniform distribution.

1 Introduction

Let $X_1, \ldots, X_n$ be a random sample from a df F and let $M_n = \max\{X_1, \ldots, X_n\}$. Then it is well known that, after appropriate normalization, the limiting distribution of M_n is one of three types. For a recent reference see DE HAAN [4] or GALAMBOS [3]. The question naturally arises as to the rate of such convergence. We will focus here on a very special case: if $F(x) = 1 - e^{-\alpha x}$, $0 \leq x < \infty$, then

$$P(M_n - \alpha^{-1}\log n \leq t) = (1 - n^{-1}e^{-\alpha t})^n 1_{[-\alpha^{-1}\log n, \infty)}(t) \to \exp(-e^{-\alpha t}) \equiv \Lambda_\alpha(t) \qquad (1)$$

as $n \to \infty$. It is apparent that a uniform rate of convergence of $(1-n^{-1}x)^n$ to e^{-x} would immediately provide a uniform rate of convergence in (1). Also note that $(1-n^{-1}x)^n \to e^{-x}$ has an obvious probability interpretation as well: if $U_1, \ldots, U_n$ is a sample of Uniform (0,1) rv's and $U_{n1} = \min\{U_1, \ldots, U_n\}$, then

$$P(nU_{nl} \geq x) = (1-n^{-1}x)^n 1_{[0,n]}(x) \to e^{-x} = P(Y \geq x) \qquad (2)$$

as $n \to \infty$ where $Y \sim$ exponential (1).

In Section 2 of this note an extremely tight bound for the maximum difference between the left and right sides of (2) is given. In Section 3 we apply the bound to obtain a uniform rate of convergence in (1), compare our special uniform bound to the more general pointwise bounds of GALAMBOS [3], and mention some related work.

2 The uniform rate of convergence of $(1-n^{-1}x)^n$ to e^{-x}

Let $\Delta_n(x) \equiv |(1-n^{-1}x)^n 1_{[0,n]}(x) - e^{-x}|$, $\Delta_n \equiv \sup_{0 \leq x < \infty} \Delta_n(x)$, and let x_n be such that $\Delta_n = \Delta_n(x_n)$ ($1_A(x) = 1$ if $x \in A$, otherwise 0). It is well-known that $\Delta_n \to 0$ as $n \to \infty$. An inequality on page 242 of WHITTAKER and WATSON [6] leads to the uniform rate

* Research supported by U.S. Army Research Office (Hall) and the National Science Foundation (Wellner).

** University of Rochester.

$n\Delta_n \leqslant 4e^{-2} \simeq .54$; Lemma 1 of AILAM [1] gives $n\Delta_n \leqslant e^{-1} \simeq .37$. We strengthen this substantially (for $n > 1$; see table below):

PROPOSITION 1. $2e^{-2} < n\Delta_n < (2+n^{-1})e^{-2}$ for all $n \geqslant 1$.

In fact (proof omitted here), the convergence of $n\Delta_n$ to $2e^{-2}$ may be shown to be monotone in n. We also find:

PROPOSITION 2. $2-n^{-1} \leqslant x_n < 2$ for all $n \geqslant 1$ (strict for $n > 1$).

PROOFS. Set $\theta \equiv x/n$ so that $0 \leqslant \theta \leqslant 1$ for $0 \leqslant x \leqslant n$. Define, for $n \geqslant 1$ and $0 \leqslant \theta \leqslant 1$,

$$f_n(\theta) \equiv e^{-n\theta} - (1-\theta)^n \geqslant 0.$$

Then we find $\Delta_n = \max_{0 \leqslant \theta \leqslant 1} f_n(\theta)$ (since $e^{-n} = f_n(1)$). When $n = 1$ the maximum of $f_n(\theta)$ is achieved at $\theta_1 = 1$ and equals e^{-1}. Hence $\Delta_1 = e^{-1}$. For $n \geqslant 2$ we proceed to find the maximum value of $f_n(\theta)$ and the maximizing θ-value, say θ_n. Since $f_n'(\theta)$ is non-positive at 0 and 1 and has a unique root in (0,1), the root is θ_n. The equation $f_n'(\theta_n) = 0$ is equivalent to

$$e^{-n\theta_n} = (1-\theta_n)^{n-1}, \tag{3}$$

and therefore $f_n(\theta_n) = \theta_n e^{-n\theta_n} > f_n(1) = e^{-n}$. Writing $h(x) \equiv xe^{-x}$, we therefore have

$$n\Delta_n = nf_n(\theta_n) = n\theta_n e^{-n\theta_n} = h(n\theta_n). \tag{4}$$

To bound $n\theta_n$, assume Proposition 2: $2-n^{-1} < n\theta_n < 2$ $(n > 1)$. This together with (4) and the fact that $h(x) \equiv xe^{-x} > 2e^{-2}$ for $1 \leqslant x < 2$ implies that $h(n\theta_n) > 2e^{-2}$. An expansion of $h(x)$ about $x = 2$ yields, with $n\theta_n \leqslant \alpha_n \leqslant 2$,

$$n\Delta_n = h(n\theta_n) = h(2) - h'(\alpha_n)(2-n\theta_n) \leqslant 2e^{-2} + e^{-2}(2-n\theta_n) < (2+n^{-1})e^{-2}.$$

Hence only Proposition 2 remains to be proved.

Proposition 2 may be verified numerically for $n = 1, \ldots, 6$ (see table below). Hereafter assume $n > 6$. Since

$$-1 - x^{-1}\log(1-x) \leqslant \tfrac{1}{2}x + \tfrac{1}{3}x^2(1-x)^{-1}$$

for $0 < x < 1$, it follows from (3) that

$$(n-1)^{-1} \leqslant \tfrac{1}{2}\theta_n + \tfrac{1}{3}\theta_n^2(1-\theta_n)^{-1}$$

which implies that

$$\theta_n \geqslant \{3(n+1)/[2(n-1)]\}\{1-(1-8(n-1)/[3(n+1)^2])^{\frac{1}{2}}\}.$$

Now $1-(1-u)^{\frac{1}{2}} \geqslant \frac{1}{2}u(1+\frac{1}{4}u)$, $0 < u < 1$, so that

$$\begin{aligned}\theta_n &\geqslant \{3(n+1)/[2(n-1)]\}\{4(n-1)/[3(n+1)^2]\}\{1+2(n-1)/[3(n+1)^2]\} \\ &= 2(n+1)^{-1}\{1+[(n-1)/(2n^2)]+\varepsilon_n\} > 2(n+1)^{-1}\{1+(n-1)/(2n^2)\} = n^{-1}(2-n^{-1})\end{aligned}$$

since $\varepsilon_n \equiv \{2(n-1)/[3(n+1)^2]\} - \{(n-1)/(2n^2)\} > 0$ for $n > 6$. This completes the proof of the left inequality of Proposition 2.

To prove the right inequality, assume it false (for some $n \geqslant 3$). Then $\theta_n \geqslant 2/n$, and since the function g defined by

$$g(x) = -1 - x^{-1} \log(1-x) = \tfrac{1}{2}x + \tfrac{1}{3}x^2 + \ldots$$

is strictly increasing in $0 < x < 1$, it follows that

$$g(\theta_n) \geqslant g(2/n) = \sum_{k=1}^{\infty} (k+1)^{-1}(2/n)^k > \sum_{k=1}^{\infty} n^{-k}.$$

On the other hand, from (3)

$$g(\theta_n) = (n-1)^{-1} = \sum_{k=1}^{\infty} n^{-k},$$

a contradiction.

The following numerical values illustrate the propositions:

n	x_n	$n\Delta_n$	$(2+n^{-1})e^{-2}$
1	1.000	*.367879 $\simeq e^{-1}$	.406
2	1.594	.324	.338
3	1.748	.304	.316
4	1.818	.295	.305
5	1.857	.290	.298
6	1.882	.287	.293
10	1.931	.280	.284
20	1.966	.275	.277
50	1.987	.272	.273
100	1.993	.272	.272
∞	2.	.270671 $\simeq 2e^{-2}$ $\simeq$	.270671

* AILAM's [1] bound on $n\Delta_n$ for all n.

3 A uniform convergence rate for the maximum of exponential rv's

Suppose that $X_1, \ldots, X_n$ is a sample from the exponential df $F(x) = 1 - e^{-\alpha x}$, $0 \leqslant x < \infty$ and $M_n = \max\{X_1, \ldots, X_n\}$. The following proposition follows immediately from Proposition 1:

PROPOSITION 3.

$$\sup_t |P(M_n - \alpha^{-1} \log n \leqslant t) - \Lambda_\alpha(t)| < n^{-1}(2+n^{-1})e^{-2}.$$

In contrast to the rapid rate of convergence given by Proposition 3, it is well-known that the distribution of the maximum of a sample of normal random variables converges to its asymptotic (double exponential) form extremely slowly; see FISHER and TIPPETT [2] or GALAMBOS [3], page 117.

GALAMBOS [3] has given a general theorem concerning the *pointwise* (in t) rate of convergence of the distributions of maxima to extreme value distributions (see Theorem 2.10.1 on page 113). In the special case considered here (with $\alpha = 1$ for simplicity) GALAMBOS' theorem yields (Examples 2.10.1, page 115)

$$|P(M_n - \log n \leqslant t) - \Lambda_1(t)| \leqslant \Lambda_1(t)[2n^{-1}e^{-2t} + 2n^{-2}e^{-4t}(1-q)^{-1}]$$

for fixed t with $(2/3)e^{-2t} \leqslant q < 1$. The uniform rate of convergence given by Proposition 3 does not seem to be implied by this latter pointwise bound. It would be of interest to have a uniform version of GALAMBOS' general pointwise bounds.

A referee has remarked that Proposition 3 might be used as a first step to obtain rates in the general case: If $F^n(a_n x + b_n) \to G(x)$ weakly, then, with $\bar{F} \equiv 1 - F$,

$$\begin{aligned}\sup_x |F^n(a_n x + b_n) - G(x)| &\leqslant \sup_x |F^n(a_n x + b_n) - e^{-n\bar{F}(a_n x + b_n)}| + \sup_x |e^{-n\bar{F}(a_n x + b_n)} - G(x)| \\ &= \sup_t |(1 - n^{-1}e^{-t})^n - \Lambda_1(t)| + \sup_x |e^{-n\bar{F}(a_n x + b_n)} - G(x)| \\ &\leqslant n^{-1}(2 + n^{-1})e^{-2} + \sup_x |e^{-n\bar{F}(a_n x + b_n)} - G(x)|\end{aligned}$$

and this latter term may be easier to handle (cf. (145) on page 114 of GALAMBOS [3]).

If the above probabilistic context is abandoned and one considers approximating e^{-x} as closely as possible by a polynomial of degree n ($(1 - n^{-1}x)^n$ being a very particular such polynomial), then geometric rates of convergence are possible; see e.g. RAHMAN and SCHMEISSER [5].

References

[1] AILAM, G., On probability properties of random sets and the asymptotic behavior of empirical distribution functions. J. Appl. Prob. **5** (1968), 196–202.

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

[3] GALAMBOS, J., The Asymptotic Theory of Extreme Order Statistics. Wiley, New York, 1978.

[4] HAAN, L. DE, On regular variation and its application to the weak convergence of sample extremes. Mathematical Centre Tract 32, Mathematisch Centrum, Amsterdam, 1975.

[5] RAHMAN, Q. I. and G. SCHMEISSER, Rational approximation to e^{-x}. II. Trans. Amer. Math. Soc. **235** (1978), 395–402.

[6] WHITTAKER, E. T. and G. N. WATSON. A Course of Modern Analysis. Fourth edition. Cambridge University Press, London, 1969.

Hall, W. J.; Wellner, Jon A. 80k:60029

The rate of convergence in law of the maximum of an exponential sample.

Statist. Neerlandica 33 (1979), *no.* 3, 151–154.

Let $X_1, X_2, \cdots, X_n$ be independent and identically distributed random variables with distribution function $F(x)$. Set $Z_n = \max(X_1, X_2, \cdots, X_n)$. Let a_n and b_n, $b_n > 0$, be constants such that $(Z_n - a_n)/b_n$ converges in distribution. The first estimate on the speed of this convergence was obtained by the reviewer in his book [*The asymptotic theory of extreme order statistics*, Wiley, New York, 1978; MR 80b:60040]. The reviewer's estimate is for pointwise convergence, which the authors extend here to uniform estimates for the case when $F(x) = 1 - e^{-ax}$, $x \geqslant 0$ $(a > 0)$.

J. Galambos (Philadelphia, Pa.)