Notes on Convergence in Law of Maxima

Jon A. Wellner

June 1, 2015

1. Introduction

For the basics of convergence in distribution of maxima of independent random variables, see van der Vaart (1998), section 21.4, pages 312-314, and Ferguson (1996), chapter 14, pages 94-100. For further more recent results in this vein, see Engelke et al. (2015) and Kabluchko (2011a, 2011b, 2014).

2. Example 1: maxima of i.i.d. standard normal

Suppose that $X_1, \ldots, X_n$ are i.i.d. with d.f. $F = \Phi$, the standard normal distribution. Then as discussed by van der Vaart (1998), section 21.4, pages 312-314, $X_{(n)} \equiv M_n \equiv \max_{1 \leq i \leq n} X_i$ satisfies

$$G_n \equiv b_n(M_n - a_n) \to_d G \sim Ev$$

where $Ev(x) = \exp(-\exp(-x))$,

$$b_n \equiv \sqrt{2 \log n},$$
$$a_n \equiv \sqrt{2 \log n} - \frac{1}{2} \frac{\log \log n + \log(4\pi)}{\sqrt{2 \log n}}$$

Moreover, the density $f_{G_n}(x) \to Ev'(x)$ and $d_{TV}(P_{G_n}, P_G) \to 0$. Figure 1 illustrates the convergence in (2.1) and the claimed convergence of densities is illustrated in Figure 2. The rate of convergence in (2.1) is $O(1/\log n)$; see Hall (1979) and Resnick (1987), page 121.

3. Example 2: maxima of i.i.d. standard exponentials

Now suppose that $X_1, \ldots, X_n$ are i.i.d. with distribution function $F$ given by $1 - F(x) = \exp(-x)$. Again let $M_n \equiv X_{(n)}$. Then it is easily seen that

$$G_n \equiv M_n - \log n \to_d G \sim Ev$$

(3.1)
This follows since

\[ P(M_n - \log n \leq x) = P(X_{(n)} \leq x + \log n) = (1 - \exp(-(x + \log n)))^n \]
\[
\left(1 - \frac{e^{-x}}{n}\right)^n \to \exp(-\exp(-x)) = Ev(x)
\]
for every \(x\). Furthermore,
\[
f_{G_n}(x) = n \left(1 - \exp(-(x + \log n))\right)^{n-1} \exp(-(x + \log n)) \\
\to \exp(-\exp(-x)) \exp(-x) = Ev'(x)
\]
for every \(x\), and hence by Scheffé’s theorem, \(d_{TV}(P_{G_n}, P_G) \to 0\). Figure 3 illustrates the convergence in (3.1) and the claimed convergence of densities is illustrated in Figure 4. The rate of convergence in both cases is \(n^{-1}\); see Hall & W (1979).

4. Example 3: supremum of a standard kernel estimator

Let \(\hat{f}_n\) be the kernel estimator of a density \(f\) on \([0, 1]\) based on a kernel \(w\) and the bandwidth \(h_n = n^{-\delta}\) with \(1/5 < \delta < 1/2\). Bickel and Rosenblatt (1973) show (under hypotheses specified in their paper) that
\[
\tilde{M}_n = \sup_{0 < t < 1} \frac{\sqrt{nh_n}|\hat{f}_n(t) - f(t)|}{\sqrt{f(t)}}
\]
satisfies the following extreme value convergence:
\[
\sqrt{2\delta \log n} \left(\frac{\tilde{M}_n}{\lambda(w)} - d_n\right) \to d Ev^2
\]
where $E v^2(x) = \exp(-2 \exp(-x))$ and where

\[
\lambda(w) \equiv \int w^2(t) dt,
\]

\[
K_1(w) \equiv \frac{w^2(A) + w^2(-A)}{2\lambda(w)},
\]

\[
K_2(w) \equiv \frac{1}{2\lambda(w)} \int \{w'(t)\}^2 dt,
\]

\[
d_n = \begin{cases} 
(2\delta \log n)^{1/2} + \frac{1}{(2\delta \log n)^{1/2}} \left\{ \frac{K_1(w)}{\sqrt{\pi}} - \frac{1}{2} \log(\delta \log n) \right\}, & \text{if } K_1(w) > 0, \\
\sqrt{2\delta \log n} + \frac{\log[K_2(w)/(2\pi)]}{\sqrt{2\delta \log n}}, & \text{otherwise}.
\end{cases}
\]

5. Example 4.

Let $G_n$ be the empirical distribution function of i.i.d. uniform(0,1) random variables and let $Z_n(t) \equiv \sqrt{n}(G_n(t) - t)/\sqrt{t(1-t)}$ for $0 < t < 1$. Then Jaeschke and Eicker showed that $\|Z_n\|_\infty \equiv \sup_{0<t<1} |Z_n(t)|$ satisfies

\[
b_n \left( \|Z_n\|_\infty - c_n/b_n \right) \to_d E v^4
\]

where

\[
b_n \equiv \sqrt{2 \log \log n}, \quad c_n \equiv 2 \log \log n + 2^{-1} (\log \log \log n - \log(4\pi)).
\]

Here $E v^4(x) = \exp(-4 \exp(-x))$. See Shorack & W (1986), page 600.


Department of Statistics  
University of Washington  
P.O. Box 354322  
Seattle, Washington 98195-4322  
U.S.A.  
e-mail: jaw@stat.washington.edu