Limit laws for normalized maxima

Def
Let $X_1, X_2, \dots$ be an iid. sequence of random variables.
Let $M_n = \max(X_1, X_2, \dots, X_n)$ denote its running maximum.
Let $F(x) = \mathbb{P}(X < x)$ be their distribution function.
Then let $x_F = \sup\{x \in \mathbb{R} : F(x) < 1\}$ denote the right endpoint of their distribution which might be $+\infty$.
Let $\overline{F}(x) = 1 - F(x) = \mathbb{P}(X \geq x)$ denote the tail probability.

Def
Let $Y_1, Y_2, \dots, Y$ r.v.'s on the same probability space. Y_n converges to Y in probability (notation: $Y_n \xrightarrow{\mathbb{P}} Y$) if $\forall \varepsilon > 0 : \mathbb{P}(|Y_n - Y| > \varepsilon) \to 0$ as $n \to \infty$.

Prop
① If $Y_n \xrightarrow{\mathbb{P}} Y$, then $Y_n \xRightarrow{d} Y$.
② If c is a constant and $Y_n \xRightarrow{d} c$, then $Y_n \xrightarrow{\mathbb{P}} c$.
In general, $Y_n \xRightarrow{d} Y$ does not imply convergence in probability, since they might not even be defined on the same probability space.

Observation
If $x < x_F$, then $\mathbb{P}(M_n < x) = F^n(x) \to 0$ as $n \to \infty$
If $x \geq x_F$, then $\mathbb{P}(M_n < x) = 1$.
Hence $M_n \xRightarrow{\mathbb{P}} x_F$ as $n \to \infty$
Since the sequence M_n is increasing in n, $M_n \to x_F$ a.s. (almost surely), i.e. with probability 1: $\mathbb{P}(M_n \to x_F) = 1$.
What happens if we allow rescaling?

Def
The random variable H has an extreme value distribution if it appears as a limit $\dfrac{M_n - b_n}{a_n} \xRightarrow{d} H$ for some $a_n > 0$ and $b_n \in \mathbb{R}$ and H is non-degenerate.

Questions

- What are the extreme value distributions?
- Under what conditions on the distribution F do we get them?
- How the constants a_n, b_n have to be chosen?

Equivalently, what are the conditions for the following probability to converge:

$$\mathbb{P}\left(\frac{M_n - b_n}{a_n} < x\right) = \mathbb{P}(M_n < a_n x + b_n) = \mathbb{P}(M_n < u_n) \text{ where } u_n = a_n x + b_n?$$

Prop (Poisson approximation)

For a given $\tau \in [0,\infty)$ and real sequence (u_n), the following are equivalent:

① $n\bar{F}(u_n) \to \tau$

② $\mathbb{P}(M_n < u_n) \to e^{-\tau}$

Proof

①$\Rightarrow$②: $\mathbb{P}(M_n < u_n) = F^n(u_n) = (1-\bar{F}(u_n))^n = \left(1 - \frac{\tau}{n} + o\left(\frac{1}{n}\right)\right)^n \to e^{-\tau}$.

②$\Rightarrow$①: ② implies $n\log(1-\bar{F}(u_n)) \to -\tau$, hence $\bar{F}(u_n) \to 0$ must hold. By Taylor expansion $n(\bar{F}(u_n) + o(\bar{F}(u_n))) \to \tau$. $\square$

Remark (probabilistic proof for ①$\Rightarrow$② for $\tau \in (0,\infty)$)

Let $B_n = \sum_{i=1}^n \mathbb{1}(X_i \geq u_n) \sim \mathrm{BIN}(n, \bar{F}(u_n))$. Then $B_n \xrightarrow{d} \mathrm{POI}(\tau)$ if and only if ① holds. On the other hand, ② is equivalent to $\mathbb{P}(B_n = 0) = \mathbb{P}(M_n < u_n) \to e^{-\tau}$.

Prop (technical)

Let $\tau \in (0,\infty)$. There is a real sequence (u_n) satisfying $n\bar{F}(u_n) \to \tau$ if and only if $\bar{F}(x_F) = 0$ and

$$\lim_{x \nearrow x_F} \frac{\bar{F}(x+)}{\bar{F}(x)} = 1.$$

Remark

Heuristically the condition above requires that the distribution does not have large atoms as we approach x_F.
Equivalently, the condition is $\lim_{x \nearrow x_F} \frac{p(x)}{\bar{F}(x)} = 0$ where $p(x) = \bar{F}(x) - \bar{F}(x+)$ is the weight of the atom at x.

Examples

The following discrete distributions do not have limit laws for normalized maxima by the proposition above.

1. $BIN(m,p)$ since $\bar{F}(x_F) > 0$.
2. $Poi(\lambda)$: for k integer $\bar{F}(k) = \sum_{j=k}^{\infty} e^{-\lambda}\frac{\lambda^j}{j!}$ $\bar{F}(k^+) = \bar{F}(k+1) = \sum_{j=k+1}^{\infty} e^{-\lambda}\frac{\lambda^j}{j!}$

$$\lim_{k\to\infty} \frac{\bar{F}(k+1)}{\bar{F}(k)} = 0 \quad \to \text{See exercises.}$$

3. $NBIN(m,p)$: for k integer $\bar{F}(k) = \sum_{j=k}^{\infty}\binom{j+m-1}{j} p^m (1-p)^j$

$$\lim_{k\to\infty} \frac{\bar{F}(k+1)}{\bar{F}(k)} = 1-p \quad \to \text{See exercises.}$$

4. $GEO(p)$: for k integer $\bar{F}(k) = \sum_{j=k}^{\infty} p(1-p)^j = (1-p)^{k-1}$

$$\frac{\bar{F}(k+1)}{\bar{F}(k)} = 1-p \quad \Rightarrow \text{no limit law.}$$

Max-stable laws

Def

The random variable H has a <u>max-stable law</u>, if there are $a_n > 0$ and $b_n \in \mathbb{R}$ such that for $X_1, X_2, \dots$ iid. with the distribution of H and $M_n = \max(X_1, \dots, X_n)$, $\frac{M_n - b_n}{a_n} \overset{d}{=} H$.

Thm

H has a max-stable law if and only if it has an extreme value distribution.

Proof

$\boxed{\Rightarrow}$ Take $X_1, X_2, \dots$ have the same distribution as H.

$\boxed{\Leftarrow}$ Fix an integer k and let $M_{nk} = \max(M_n^{(1)}, \dots, M_n^{(k)})$ where $M_n^{(j)} = \max(X_{(j-1)n+1}, \dots, X_{jn})$ for which $\frac{M_n^{(j)} - b_n}{a_n} \overset{d}{\Rightarrow} H$

Hence $\frac{M_{nk} - b_n}{a_n} = \max\left(\frac{M_n^{(1)} - b_n}{a_n}, \dots, \frac{M_n^{(k)} - b_n}{a_n}\right) \overset{d}{\Rightarrow} \max(H_1, \dots, H_k)$ where $H_1, \dots, H_k$ are iid with the distr. of H

On the other hand $\frac{M_{nk} - b_{nk}}{a_{nk}} \overset{d}{\Rightarrow} H$

By convergence of types: $\frac{\max(H_1, \dots, H_k) - B_k}{A_k} \overset{d}{=} H$ where $A_k = \lim_{n\to\infty} \frac{a_{nk}}{a_n}$, $B_k = \lim_{n\to\infty}\frac{b_{nk} - b_n}{a_n}$. □