ICFP M2 - Statistical physics 2 – TD n° 1
Extreme values distributions

Grégory Schehr, Francesco Zamponi

We consider a sequence X_1, X_2, ... of independent identically distributed (i.i.d.) random variables, with a common distribution equal to the one of X. One is often interested in the behavior as $n \to \infty$ of:

- the maximum $M_n = \max(X_1, \dots, X_n)$, and
- the sum $S_n = X_1 + \cdots + X_n$

of n such random variables. In both cases there only exists a few "universality classes" for the possible limit behaviors, independently on most of the "microscopic" details of the law of X. The lecture has covered the case of the sum (with stable random variables generalizing the central limit theorem), we consider in this problem the behavior of the maximum M_n.

We will denote R_X the right-edge of the support of X, defined as $R_X = \inf\{x : F_X(x) = 1\}$, which can be equal to $+\infty$.

One says that a random variable Z is degenerate if it takes a constant value c almost surely, then $F_Z(z) = 0$ for $z < c$ and $F_Z(z) = 1$ for $z \geq c$.

1 Extreme value distributions

1. Explain why $F_{M_n}(x) = (F_X(x))^n$.
2. Draw the shape of $F_{M_n}(x)$ as n grows, distinguishing the cases:
   (a) $R_X = +\infty$,
   (b) $R_X < \infty$ and F_X is continuous in R_X,
   (c) $R_X < \infty$ and F_X is discontinuous in R_X.
3. Show that $M_n \stackrel{\mathrm{d}}{\to} R_X$ as $n \to \infty$.
4. In order to have a more precise description of the behavior of M_n one needs to shift and rescale it, we thus define $\widehat{M}_n = \frac{M_n - a_n}{b_n}$, where a_n and $b_n > 0$ are two sequences chosen appropriately so that $\widehat{M}_n$ converges in distribution to a non degenerate random variable (if possible) when $n \to \infty$.
   (a) Express the distribution function $F_{\widehat{M}_n}$ of the rescaled maximum in terms of F_X.
   (b) Show that if a_n and b_n are chosen in such a way that $F_X(a_n + \widehat{x}\, b_n) = 1 - \frac{\gamma(\widehat{x})}{n} + o\left(\frac{1}{n}\right)$, where $\widehat{x}$ and $\gamma(\widehat{x})$ are finite when $n \to \infty$, then $\widehat{M}_n$ has indeed a non-trivial limit, and express its distribution function.
5. Consider first that X has an exponential distribution of parameter 1, i.e. $F_X(x) = 1 - e^{-x}$ for $x \geq 0$, $F_X(x) = 0$ for $x \leq 0$. Find the values of a_n and b_n that ensures the convergence of $\widehat{M}_n$ towards a random variable whose distribution function is

$$G_{\mathrm{G}}(x) = e^{-e^{-x}} \ . \tag{1}$$

   Draw the shape of this distribution function, and of the associated density.

6. Same questions when X takes values between 0 and 1, with a distribution function on this interval given by $F_X(x) = 1-(1-x)^\alpha$, with $\alpha>0$ a parameter; the limit distribution function for $\widehat{M}_n$ is now
$$G_{\mathrm{W}}(x) = \begin{cases} e^{-(-x)^\alpha} & \text{for} \quad x\leq 0 \\ 1 & \text{for} \quad x \geq 0 \end{cases} . \tag{2}$$

7. Same questions when X is a Pareto random variable, taking values on $[1,\infty[$, with a distribution function $F_X(x) = 1-x^{-\alpha}$, with again $\alpha>0$, the limit distribution being
$$G_{\mathrm{F}}(x) = \begin{cases} 0 & \text{for} \quad x\leq 0 \\ e^{-x^{-\alpha}} & \text{for} \quad x > 0 \end{cases} . \tag{3}$$

8. Consider finally the case where X admits the density $f_X(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$, i.e. X is a standard Gaussian random variable. Check that with the choice
$$a_n = \sqrt{2\log n} - \frac{\log\log n + \log 4\pi}{2\sqrt{2\log n}} \ , \qquad b_n = \frac{1}{\sqrt{2\log n}} \ , \tag{4}$$
the normalized maximum $\widehat{M}_n$ converges to the distribution given in (1). *Indication :* the distribution function of Gaussian random variables admits the asymptotic expansion
$$F_X(x) = 1 - \frac{1}{x}\frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} + O\left(\frac{e^{-\frac{x^2}{2}}}{x^3} \right) \qquad \text{as } x\to\infty \ . \tag{5}$$

It turns out that these three distributions, named respectively after Gumbel, Weibull and Fréchet, are the only possible non degenerate limits for $\widehat{M}_n$, modulo a shift and a rescaling, whatever the original distribution of X. Note that for some X there is no choice of a_n and b_n that allows for a non degenerate limit (this is the case of the example of question 2c, but also of Poisson random variables).

The type of the distribution function for the limit of $\widehat{M}_n$ depends on the behavior of F_X at the edge of its support R_X:

- If R_X is finite and F_X reaches 1 with a power-law behavior (possibly multiplied by a slowly varying function) then X is in the domain of attraction of the Weibull law.
- The Fréchet law is instead observed if R_X is infinite and F_X reaches 1 with an inverse power-law behavior (possibly multiplied by a slowly varying function).
- If F_X reaches 1 faster than any power-law at R_X (be this edge finite or not) then the Gumbel distribution will apply.

2 (Optional) Proof of the three types theorem

We sketch here the proof of the statement made above: the only possible limit distributions for $\widehat{M}_n$ are the Gumbel, Weibull and Fréchet laws.

We will admit the following intuitive result. Suppose that $(X_n)_{n\geq 1}$ is a sequence of random variables, with two rescalings allowing for a non-trivial limit. Hence, for large n,

$$X_n \sim a_n + b_n Y \sim c_n + d_n Z \ , \tag{6}$$

with Y and Z being of order one when $n \to \infty$. If both Y and Z must stay finite, it follows that

$$Z \sim \frac{a_n - c_n}{d_n} + \frac{b_n}{d_n} Y \sim A + BY \ , \tag{7}$$

with

$$\frac{b_n}{d_n} \to B \ , \qquad \frac{a_n - c_n}{d_n} \to A \ . \tag{8}$$

We can choose the sign of Y and Z in such a way that $B > 0$.

The proof then proceeds as follows (*you are not required to be fully rigorous*):

1. Suppose that a_n and b_n are chosen such that $\widehat{M}_n$ converges towards a random variable Y with a distribution function G. Show that for any positive integer m there exist A_m and $B_m > 0$ such that
$$\max(Y_1, \dots Y_m) \overset{\mathrm{d}}{=} \frac{Y - A_m}{B_m} \quad \text{i.e.} \quad G^m(x) = G(A_m + B_m x) \tag{9}$$
where $Y_1, \dots, Y_m$ are independent random variables with the same distributions as Y; Y is thus said to be stable under the max operation. Express A_m and B_m in terms of a_n and b_n.
Hint: write $M_n \sim a_n + b_n Y$ for large n, consider the random variable $Z = \max(Y_1, \cdots, Y_m)$, and reason as in Eq. (7).

2. Generalize this result to
$$G^s(x) = G(A(s) + B(s)x) \tag{10}$$
for all reals $s > 0$.
Hint: approximate $s \sim p/q$ for integers p, q. Use Eq. (9) and its inverse, $G(x)^{1/m} = G\left(\frac{x - A_m}{B_m}\right)$.

3. Show that the functions $A(s)$ and $B(s)$ are solutions of
$$\begin{cases} B(st) = B(s)B(t) \ , \\ A(st) = A(s) + B(s)A(t) = A(t) + B(t)A(s) \ , \end{cases} \tag{11}$$
for all $s, t > 0$.

4. In the following we assume for simplicity that A and B are differentiable. Show that $B(s) = s^\theta$, where θ is an arbitrary real parameter.

5. Show that if $\theta = 0$ the distribution function G is of the Gumbel form (modulo an affine change of variables).

6. Assuming now $\theta > 0$, prove that G is of the Weibull type with $\alpha = 1/\theta$. Similarly the Fréchet distribution is obtained when $\theta < 0$.