ICFP M2 - Statistical physics 2 – TD n° 1
Extreme values distributions
Solution of the last exercise

Grégory Schehr, Guilhem Semerjian

January 2020

1. Intuitively, the statement is

$$X_n \approx a_n + b_n Y \approx c_n + d_n Z \Rightarrow Z \approx \frac{a_n - c_n}{d_n} + \frac{b_n}{d_n} Y \Rightarrow \lim_{n\to\infty} \frac{a_n - c_n}{d_n} = A \qquad \lim_{n\to\infty} \frac{b_n}{d_n} = B \ . \quad (1)$$

The precise translation of the statement in terms of distribution functions is : if $F_{X_n}(a_n + b_n x) \to G(x)$ and $F_{X_n}(c_n + d_n x) \to H(x)$ with G and H non-trivial distribution functions, then

$$\frac{b_n}{d_n} \to B \in]0,\infty[\ , \qquad \frac{a_n - c_n}{d_n} \to A \ \text{ with } A \text{ finite and } \ G(x) = H(A + Bx) \ . \quad (2)$$

2. We choose a_n and b_n such that

$$F_{M_n}(a_n + b_n x) = (F_X(a_n + b_n x))^n \underset{n\to\infty}{\to} G(x) \ . \quad (3)$$

Fixing a positive integer m, a subsequence of this limit yields

$$F_{M_{mn}}(a_{mn} + b_{mn} x) \underset{n\to\infty}{\to} G(x) \ . \quad (4)$$

On the other hand,

$$F_{M_{mn}}(a_{mn} + b_{mn} x) = (F_{M_n}(a_{mn} + b_{mn} x))^m \ , \quad (5)$$

hence

$$F_{M_n}(c_n + d_n x) \underset{n\to\infty}{\to} H(x) = G(x)^{\frac{1}{m}} \quad \text{with} \quad c_n = a_{mn} \ , \quad d_n = b_{mn} \ . \quad (6)$$

From the statement of the previous question, this implies the existence of $A(m)$ and $B(m) > 0$ such that

$$G(x) = H(A(m) + B(m)x) = G(A(m) + B(m)x)^{\frac{1}{m}} \ , \qquad \text{i.e.} \quad G^m(x) = G(A(m) + B(m)x) \ . \quad (7)$$

To generalize this to real s, we notice that (using $\lfloor u \rfloor$ for the integer part of the real u)

$$F_{M_{\lfloor ns \rfloor}}(a_{\lfloor ns \rfloor} + b_{\lfloor ns \rfloor} x) \underset{n\to\infty}{\to} G(x) \qquad \text{and} \quad F_{M_{\lfloor ns \rfloor}}(a_{\lfloor ns \rfloor} + b_{\lfloor ns \rfloor} x) = F_{M_n}(a_{\lfloor ns \rfloor} + b_{\lfloor ns \rfloor} x)^{\frac{\lfloor ns \rfloor}{n}} \quad (8)$$

hence with $c_n = a_{\lfloor ns \rfloor}$ and $d_n = b_{\lfloor ns \rfloor}$ one has

$$F_{M_n}(c_n + d_n x) = F_{M_{\lfloor ns \rfloor}}(a_{\lfloor ns \rfloor} + b_{\lfloor ns \rfloor} x)^{\frac{n}{\lfloor ns \rfloor}} \underset{n\to\infty}{\to} H(x) = G(x)^{\frac{1}{s}} \ , \quad (9)$$

which implies the existence of $A(s)$ and $B(s)$ with $G^s(x) = G(A(s)x + B(s))$.

3. By computing $G^{st}(x)$ in two different ways one gets

$$\begin{aligned} G^{st}(x) &= G(A(st) + B(st)x) = (G^s(x))^t = G^t(A(s) + B(s)x) & (10) \\ &= G(A(t) + B(t)(A(s) + B(s)x)) = G(A(t) + B(t)A(s) + B(t)B(s)x) \ . & (11) \end{aligned}$$

As $G(x)$ is the distribution function of a non-trivial random variable, $G(x) = G(\alpha + \beta x)$ for all x implies $\alpha = 0$ and $\beta = 1$, hence the equations satisfied by the functions A and B

$$\begin{cases} B(st) = B(s)B(t) \ , \\ A(st) = A(t) + B(t)A(s) = A(s) + B(s)A(t) \ , \end{cases} \quad (12)$$

for all $s, t > 0$, the last equality being obtained by symmetry between s and t.

4. Taking the derivative with respect to s of the first equation, then setting $t = 1$ yields $sB'(s) = B(s)B'(1)$. This implies $B(1) = 1$, and the differential equation can then be easily integrated to obtain $B(s) = s^\theta$, where θ is an arbitrary real parameter. Actually this is the only type of solution of the equation $B(st) = B(s)B(t)$ with the weaker assumption that $B(s)$ is continuous.
5. If $\theta = 0$ one has $B(s) = 1$ for all s, hence $A(s)$ is solution of the functional equation $A(st) = A(s) + A(t)$. We see that $e^{A(s)}$ is thus solution of the same functional equation than the one on $B(s)$ solved in the previous question, which implies $A(s) = -c \ln s$ with c an undetermined constant. We thus have an equation on the distribution function of the limit random variable, $G^s(x) = G(x - c \ln s)$. As the left-hand-side is a decreasing function of s one must have $c > 0$. Taking the logarithm of this equation yields $\ln G(x) = \frac{\ln G(x - c \ln s)}{s}$, for all x and $s > 0$. Choosing x_0 such that $G(x_0) = 1/e$, and s such that $x - c \ln s = x_0$, yields $G(x) = \exp[-\exp[-\frac{x-x_0}{c}]]$. We have thus proven that if $\theta = 0$ the distribution G is of the Gumbel form, modulo the affine change of variables with parameters x_0 and c.
6. If we assume now that $\theta > 0$, hence $B(s) = s^\theta$, we need to determine the function $A(s)$ from the equation $A(s) + s^\theta A(t) = A(t) + t^\theta A(s)$. Taking an arbitrary value of $t \neq 1$ we rewrite this equation as
$$A(s) = (1 - s^\theta) \frac{A(t)}{1 - t^\theta} \ . \tag{13}$$
The last fraction being independent of s, we have determined $A(s)$ modulo a multiplicative constant, to be denoted x_0. This yields $G^s(x) = G(x_0(1 - s^\theta) + s^\theta x) = G(x_0 + s^\theta(x - x_0))$. We need to constrain x to $x < x_0$ as the left-hand-side is decreasing with s. Taking the logarithm of this equation yields $\ln G(x) = \frac{\ln G(x_0(1-s^\theta)+s^\theta x)}{s}$ for all $x < x_0$ and $s > 0$. This can be solved by choosing s such that $x_0 + s^\theta(x - x_0) = x_1$ independently of x, which yields $G(x) = \exp\left[-\left(\frac{x_0 - x}{w}\right)^{\frac{1}{\theta}}\right]$ with w a constant. This is the Weibull distribution with $\alpha = 1/\theta$, up to the affine change of variables with parameters x_0 and w. The case $\theta < 0$ is treated exactly in the same way, but with now the constraint $x > x_0$.