Statistical Theory of Extremes
Homepage: http://www.gathacognition.com/book/gcb14
http://dx.doi.org/10.21523/gcb1

Part 3
Multivariate extremes

Chapter 10

Bivariate Extremes: Probabilistic Results

J. Tiago de Oliveira
Academia das Ciências de Lisboa (Lisbon Academy of Sciences), Lisbon, Portugal.

Abstract

An asymptotic distribution of bivariate extremes is studied to obtain the asymptotic probabilistic behaviour and to provide bivariate models of (asymptotic) extremes. The distribution of bivariate extremes and the limiting distribution of maxima are discussed with complementary basic, correlation and regression results. The classical correlation coefficient is linearly invariant. Orthogonal polynomials with respect to $\Lambda(x)$ is used to improve the regression lines.

Published Online

23 June 2017

Keywords

Bivariate asymptotic distributions;
Univariate extremes;
Correlation;
Regression;
Bivariate reduced extremes.

Editor(s)

J.C. Tiago de Oliveira

10.1 Introduction

Bivariate asymptotic distributions of maxima are useful for the analysis of many concrete problems such as the greatest ages of death for men and women, each year, whose distribution, naturally, splits in the product of the margins, by independence; floods at two different places on the same river, each year; bivariate extreme meteorological data (pressures, temperature, wind velocity, etc.), each week; largest waves, each week, etc. The same can be said for the study of minima. Extensions to multivariate distributions, more complex, will be made later.

Originally published in 'Statistical Analysis of Extremes', 1997, 2016
http://dx.doi.org/10.21523/gcb1.1717

Evidently, the aim of study of asymptotic distributions of bivariate extremes is to obtain the asymptotic probabilistic behaviour, and also to provide bivariate models of (asymptotic) extremes that fit observed data. But, as will see in the next chapter, only some problems have been solved and the methods found so far cover a much smaller area than the theory for univariate extremes.

10.2 The distribution of bivariate extremes

The theory of limiting bivariate (and multivariate) extremes, to a certain extent, follows the same lines of limiting univariate extremes. For simplicity of the exposé we now consider the bivariate case; the multivariate case will appear as an extension in a following chapter.

Let $(X_1, Y_1), \dots, (X_k, Y_k), \dots$ be a sequence of i.i.d. random pairs with distribution function $F(x, y) = \mathrm{Prob}\{X \le x, Y \le y\}$ and survival function $S(x, y) = \mathrm{Prob}\{X > x, Y > y\} = 1 + F(x, y) - F(x, +\infty) - F(+\infty, y)$; the inversion of the last relation expresses $F(x, y)$ in $S(x, y)$ by $F(x, y) = 1 + S(x, y) - S(x, -\infty) - S(-\infty, y)$ because, as known, we have $F(x, +\infty) + S(x, -\infty) = 1$ and $F(+\infty, y) + S(-\infty, y) = 1$. The survival function plays the same rôle for minima as the distribution function played for maxima.

The random pair $(\max_1^k X_i, \max_1^k Y_i)$ has the distribution function

$$F_k(x, y) = \mathrm{Prob}\{\max_1^k X_i \le x, \max_1^k Y_i \le y\} = F^n(x, y).$$

The probability of $(\max_1^k X_i, \max_1^k Y_i)$ being an observed point is, evidently,

$$\pi_k = k\, \mathrm{Prob}\,\{\max_1^k X_i = X_1, \max_1^k Y_i = Y_1\} = k \int \int_{-\infty}^{+\infty} F^{k-1}(x, y)\ d\,F(x, y).\ .$$

If $F(x, y)$ splits into its margins $F(x, y) = F(x, +\infty).F(+\infty, y)$, we have $\pi_k = 1/k.$ It is well known (see Fréchet, 1951) that $\max(0, F(x, +\infty) + F(+\infty, y) - 1) \le F(x, y) \le \min\,(F(x, +\infty), F(+\infty, y))$ with the bounds attained; if $F(x, y)$ is equal to the LHS we have $\pi_k = 0$, and for the RHS we get $\pi_k = 1$ because in the first case we have $F(x,+\infty) + F(+\infty,y) = 1$ with probability 1, and in the second we have $F(x,+\infty) = F(+\infty,y)$ with probability 1.

For the random pair $(\min_1^k X_i, \ \min_1^k Y_i)$ we have $S_k(x,y) = S^k(x,y)$ in an analogous way, and the distribution function of the pair of minima is

$$\text{Prob}\{\min_1^k X_i \le x, \min_1^k Y_i \le y\} = 1 + S^k(x,y) - S^k(x,-\infty) - S^k(-\infty,y) =$$

$$1 + (1 + F(x,y) - F(x,+\infty) - F(+\infty,y))^k - (1 - F(x,+\infty))^k - (1 + F(+\infty,y))^k\,.$$

The probability π'_k of $(\min_1^k X_i, \ \min_1^k Y_i)$ being an observed point is also $1/k$ for the independence case and 0 and 1 if $F(x,y)$ is equal to the LHS or RHS of the corresponding Fréchet inequality.

As before we will stick essentially to the study of maxima, the conversion to minima being immediate by the relations $\min_1^k X_i = -\max_1^k(-X_i)$ and $\min_1^k Y_i = -\max_1^k(-Y_i)$; note that the distribution function of $(-X,-Y)$ is $\text{Prob}\{-X \le x, -Y \le y\} = \text{Prob}\{X \ge -x, Y \ge -y\} = S(-x,-y)$ in the continuity points (with right-continuity continuation in the discontinuity set) and thus, here, all over the plane because the limiting distributions are continuous.

Note, as a hint for applications, if (X_i, Y_i) are floods of the same river at two different points they must be dealt with as maxima; but if (X_i, Y_i) are droughts at the same two points they must be dealt with as minima, i.e. $(-X_i, -Y_i)$ will be dealt with as maxima.

10.3 The limiting distribution of maxima

An important question, both theoretical and applied, is to find out, analogously, as was done in the univariate case, if there exist positive linear transforms of $(\max_1^k X_i, \max_1^k Y_i)$ such that the reduced random pair

$$\left((\max_1^k X_i - \lambda_k)/\delta_k, (\max_1^k Y_i - \lambda'_k)/\delta'_k\right),$$

whose distribution function is

$$\text{Prob}\{(\max_1^k X_i - \lambda_k)/\delta_k \le x, (\max_1^k Y_i - \lambda'_k)/\delta'_k \le y\} = F^k(\lambda_k + \delta_k\, x, \lambda'_k + \delta'_k\, y) \xrightarrow{w} L(x, y),$$ i.e., has a limiting, non-degenerate and proper distribution.

As a consequence we know that $F^k(\lambda_k + \delta_k\, x, +\infty) \xrightarrow{w} L(x, +\infty)$ and $F^k(+\infty, \lambda'_k + \delta'_k\, y) \xrightarrow{w} L(+\infty, y)$ — see Deheuvels (1984).

Evidently, as we dealing with maxima, the marginal distribution functions $L(x, +\infty)$ and $L(+\infty, y)$ must be of any of the three forms — Weibull, Gumbel or Fréchet. As they are easily translated from one to another, as seen before, we will suppose that the margins have a reduced Gumbel distribution function, i.e. $L(x, +\infty) = \Lambda(x), L(+\infty, y) = \Lambda(y)$. As the margins are continuous, so is $L(x, y)$ and, thus, the convergence $F^k\left(\lambda_k + \delta_k\, x, \lambda'_k + \delta'_k\, y\right) \xrightarrow{w} L(x, y)$ is also uniform. For notational convenience, we will denote by $\Lambda(x, y)$ the limiting distribution functions $L(x, y)$ when the margins have a reduced Gumbel distribution function.

The special case where we have minima with reduced exponential distribution will also be considered, sometimes by duality, but the kernel of the Chapter, as well as the next ones, is the case of maxima with reduced Gumbel margins.

Let us show that $\Lambda(x, y) = e^{-(e^{-x}+e^{-y})k(y-x)}$ where $k(.)$, called the *dependence function* must satisfy some conditions to be seen later.

The proof of this result is simple.

As $F^{k}(\lambda_k + \delta_k\, x, \lambda'_k + \delta'_k\, y) \to \Lambda(x,y)$ also

$F^{km}(\lambda_{km} + \delta_{km}\, x, \lambda'_{km} + \delta'_{km}\, y) \to \Lambda(x,y)$ or

$F^{km}(\lambda_{km} + \delta_{km}\, x, \lambda'_{km} + \delta'_{km}\, y) \to \Lambda^{1/m}(x,y)$.

By Khintchine's convergence of types theorem we know that there exist coefficients

$\alpha_m, \beta_m > 0, \alpha'_m, \beta'_m > 0$ such that

$\Lambda^{1/m}(x,y) = \Lambda(\alpha_m + \beta_m\, x, \alpha'_m + \beta'_m\, y)$.

Using this relation for the margins (i.e., putting $y = +\infty$ or $x = +\infty$) we get $\Lambda(\alpha_m + \beta_m\, x) = \Lambda^{1/m}(x), \Lambda(\alpha'_m + \beta'_m\, y) = \Lambda^{1/m}(y)$ and thus $\alpha_m = \alpha'_m = \log m, \beta_m = \beta'_m = 1$, and so

$\Lambda(x,y) = \Lambda^{m}(\log m + x, \log m + y)$.

Let us take $m = [k\,t], t > 0$, the integer part of $(k\,t)$. We have

$\Lambda(x,y) = \Lambda^{[k\,t]}(\log\,[k\,t] + x, \log\,[k\,t] + y)$.

Substituting $x - \log k$ for x and $y - \log k$ for y, we get

$\Lambda(x - \log k, y - \log k) = \Lambda^{k}(x,y)$

and also $\Lambda(x - \log k, y - \log k) = \Lambda^{[k\,t]}(\log\,[k\,t] - \log k + x\,, \log\,[k\,t] - \log k + y)$

so we have $\Lambda(x,y)^{k/[k\,t]} = \Lambda(\log[k\,t]/k + x, \log\,[k\,t]/k + y)$ and

letting $k \to \infty$, by continuity, we get the final stability relation $(t > 0)$:

$\Lambda(x,y) = \Lambda^{t}(\log t + x, \log t + y)$.

For $x + \log t = 0$ it gives

$\Lambda(x,y) = [\Lambda(0, y - x)]^{e^{-x}}$,

which can be put under the form

$\Lambda(x,y) = \exp\{-(e^{-x} + e^{-y})\, k(y - x)\}$

as said before. The dependence function expresses the probabilistic inter-relation, or association, between the margins.

An actual characterization of the distribution function $\Lambda(x, y)$ can be made in the following way. It is immediate that a random pair (X, Y) with distribution function $\Lambda(x, y)$ is such that $\max(X + a, Y + b)$ has a Gumbel distribution function with a location parameter. But, also, the converse is true. In fact putting $z - a = x, z - b = y$ we have $L(x, y) = \Lambda(z - \varphi(z - x, z - y))$ independent of z, from which we get $\varphi(\xi, \eta) = \xi + \varphi(0, \eta - \xi)$, and then $(1 + e^{-w})\, k(w) = e^{\varphi(0,-w)}$. Thus:

$\Lambda(x, y)$ is the only distribution function $L(x, y)$ such that $Prob\{max(X + a, Y + b) \leq z\} = L(z - a, z - b) = \Lambda(z - \varphi(a, b))$.

Another approach may be the following: let us call the *structure function* of a bivariate distribution function $F(x, y)$ with continuous margins $A(x) = F(x, +\infty)$ and $B(y) = F(+\infty, y)$ the function $\bar{S}(\xi, \eta)$ such that

$\bar{S}(A(x), B(y)) = F(x, y);$

$\bar{S}(\xi, \eta)$ is a bivariate distribution function with uniform margins, defined in the unit square $[0, 1] \times [0, 1]$ and is continuous.

Then any structure function of the limiting distribution functions of bivariate extremes verifies the functional equation $(0 < \omega < +\infty)$

$\bar{S}^{\omega}(\xi, \eta) = \bar{S}(\xi^{\omega}, \eta^{\omega});$

Let S_0 denote the (initial) structure function of the distribution function with margins $A(x)$ and $B(y)$ and suppose, as before, that the marginal limiting distributions are $\Lambda(x)$ and $\Lambda(y)$, with $\lambda_k, \delta_k > 0, \lambda'_k, \delta'_k > 0$, as a system of attraction coefficients. We then have

$\xi_k = F^k(\lambda_k + \delta_k\, x, +\infty) \to \Lambda(x)$

$\eta_k = F^k(+\infty, \lambda'_k + \delta'_k\, y) \to \Lambda(y)$

and $F^k(\lambda_k + \delta_k\, x, \lambda'_k, \delta'_k\, y) = \bar{S}_0^k(\xi_k^{1/k}, \eta_k^{1/k})$. Let us now show that $\bar{S}_0^k(\xi_k^{1/k}, \eta_k^{1/k})$ has a limit $\bar{S}(\xi, \eta)$ when, $\xi_k \to \xi, \eta_k \to \eta$ if and only if $\bar{S}_0^k\, (\xi^{1/k}, \eta^{1/k}) \to \bar{S}(\xi, \eta)$. The 'if' part is obvious; let us show the converse.

As $\xi_k \to \xi, \eta_k \to \eta$, for $\eta > N(\varepsilon)$, we have

$|\xi_k - \xi| \leq \varepsilon, |\eta_k - \eta| \leq \varepsilon$

and, thus,

$(\xi-\varepsilon)^{1/k} \le \xi_k^{1/k}(\xi+\varepsilon)^{1/k}$

$(\eta-\varepsilon)^{1/k} \le \eta_k^{1/k}(\eta+\varepsilon)^{1/k}$

so that

$\bar{S}_0^k((\xi-\varepsilon)^{1/k},(\eta-\varepsilon)^{1/k}) \le \bar{S}_0^k(\xi_k^{1/k},\eta_k^{1/k}) \le \bar{S}_0^k((\xi+\varepsilon)^{1/k},(\eta+\varepsilon)^{1/k})$,
and every possible limit of $\bar{S}_0^k(\xi_k^{1/k},\eta_k^{1/k})$ is between $\bar{S}(\xi-\varepsilon,\eta-\varepsilon)$ and $\bar{S}(\xi+\varepsilon,\eta+\varepsilon)$; owing to the continuity of $\bar{S}(\xi,\eta)$ we have the desired result.

We must now characterize $\bar{S}(\xi,\eta)$. As $\bar{S}_0^k(\xi^{1/k},\eta^{1/k}) \to \bar{S}(\xi,\eta)$ and $\bar{S}_0^{k'k}(\xi^{1/k'k},\eta^{1/k'k}) \to \bar{S}(\xi,\eta)$, putting $\xi' = \xi^{1/k'}, \eta' = \eta^{1/k'}$ we have

$\bar{S}^{k'k}\left(\xi^{1/k'k},\eta^{1/k'k}\right) = [\bar{S}^k(\xi'^{1/k},\eta'^{1/k})]^{k'}$,

and letting $k \to \infty$ we have

$\bar{S}(\xi,\eta) = \bar{S}^{k'}(\xi',\eta')$

that is $\bar{S}^{k'}(\xi',\eta') = \bar{S}(\xi'^{k'},\eta'^{k'})$, and subsequently $\bar{S}^{m/k}(\xi',\eta') = \bar{S}(\xi'^{m/k},\eta'^{m/k})$. By the continuity of $\bar{S}(\xi,\eta)$, for every $\omega > 0$, we have

$\bar{S}^{\omega}(\xi,\eta) = \bar{S}(\xi^{\omega},\eta^{\omega})$.

The solution of this functional equation can be written under the form

$\bar{S}(\xi,\eta) = (\xi\,.\,\eta)^{k(\log\frac{\log\xi}{\log\eta})}$.

Taking, as stated, the reduced margins to be $\Lambda(x) = \exp(-e^{-x})$ and $\Lambda(y) = \exp(-e^{-y})$, we have the following fundamental result:

Any limiting distribution function of bivariate pairs with reduced Gumbel margins is of the form

$\Lambda(x,y) = exp\{-(e^{-x}+e^{-y})\,k(y-x)\} = \Lambda(x)\Lambda(y)\}^{k(y-x)}$, *and conversely*

When the asymptotic extremal pair $(\tilde{X},\tilde{Y})$ does not have reduced margins, that is $\tilde{X}$ and $\tilde{Y}$ have location parameters λ_x, λ_y and dispersion parameters $\delta_x > 0, \delta_y > 0$, the distribution function of $(\tilde{X},\tilde{Y})$, as $((\tilde{X}-\lambda_x)/\delta_x, (\tilde{Y}-\lambda_y)/\delta_y)$ has reduced margins, is $\Lambda((\tilde{x}-\lambda_x)/\delta_x, (\tilde{y}-\lambda_y)/\delta_y)$.

We have shown in two ways that the limiting and the stable distribution functions of extremal pairs form the same class.

Let us now deal briefly with minima as $\underset{\sim}{X}_k = \min(X_1, \dots, X_k) = -\max(-X_1, \dots, -X_k)$ and $\underset{\sim}{Y}_k = \min(Y_1, \dots, Y_k) = -\max(-Y_1, \dots, -Y_k)$, the asymptotic distribution function of minima with reduced Gumbel margins $1 - \exp(-e^x)$ and $1 - \exp(-e^y)$, if they exist, and correspondence with $\{\Lambda(x) \,.\, \Lambda(y)\}^{k(y-x)}$, is

$$\Psi(x, y) = 1 - \exp(-e^x) - \exp(-e^y) + \exp(-(e^x + e^y)\, k(x - y)).$$

Using the fact that $\min_1^k (Y_i) = -\max_1^k(-Y_i)$ we can also prove easily that the limiting distribution of $(\min_1^k X_i, \max_1^k Y_i)$, if it exists, is

$$\Theta(x, y) = \Lambda(y) - \Lambda(-x, y) = \exp(-e^{-y}) - \exp[-(e^x + e^{-y})k(y + x)].$$

We have obtained the dependence function $k(w)$ obviously continuous and non-negative for $\Lambda(x, y)$ to be a distribution function as well as the corresponding functions in other equivalent representations. These results are well known. They can be found, with different forms for the margins, in Finkelshteyn (1953), Tiago de Oliveira (1958), Geffroy (1958/59) and Sibuya (1960), with a synthesis of the results in Tiago de Oliveira (1962/63). Subsequent results are in Tiago de Oliveira (1975), (1980) and (1984). Berman (1962) gives also some limiting results and Galambos (1978) contains also a recent account.

Let us now obtain some inequalities for $\Lambda(x, y)$ or $k(w)$ which will simplify some of the derivations below. Although $\Lambda(x, y)$ is a continuous function, it does not necessarily have derivatives, and consequently we cannot expect all bivariate maxima random pairs with distribution function $\Lambda(x, y) = [\Lambda(x)\, \Lambda(y)]^{k(y-x)}$ to have a planar density. From the Boole-Fréchet inequality — Fréchet (1951) —

$$\max(0, \Lambda(x) + \Lambda(y) - 1) \le \Lambda(x, y) \le \min(\Lambda(x)\Lambda(y))$$

we have, replacing x and y by $x + \log k$ and $y + \log k$, raising to the power k and letting $k \to \infty$, the limit inequality

$\Lambda(x)\,\Lambda(y) \le \Lambda(x,y) \le \min(\Lambda(x)\,\Lambda(y))$

or

$\exp\{-(e^{-x} + e^{-y}) \le \Lambda(x,y)\} \le \exp(-e^{-\min(x,y)})$.

The LHS inequality shows that we have positive association, that is $\text{Prob}\{X \le x, Y \le y\} + \text{Prob}\{X > x, Y > y\}$ is larger or equal to the corresponding values the case of for independence for all x and y, which means that if we have a large (small) value of X, the value of Y, tends also to be large (small). $\Lambda(x).\Lambda(y)$ is, evidently, the independence situation; $\min(\Lambda(x), \Lambda(y))$ is called the *diagonal or complete dependence case* where we have, for reduce values, $\text{Prob}\{Y = X\} = 1$. As regards $k(w)$, we have

$$(1/2 \le)\ \frac{\max(1, e^w)}{1 + e} \le k(w) \le 1;$$

$k_D(w) = \frac{\max(1,e^w)}{1+e^w} = \frac{1}{1+e^{-|w|}}$ is the dependence function corresponding to the diagonal case and $k_1(w) = 1$ is the dependence function for the independence situation. Positive association, shown by Sibuya (1960), is a result that could be expected.

Remark that for $\bar{S}(\xi,\eta)$ defined above we have $\xi\eta \le \bar{S}(\xi,\eta) \le \min(\xi,\eta)$.

If there exists a planar density almost everywhere, i.e., if $k''(w)$ exists almost everywhere, $k(w)$ must satisfy the relations:

$k(-\infty) = k(+\infty) = 1,$

$[(1 + e^w)\, k(w)]' \ge 0,$

$[(1 + e^{-w})\, k(w)]' \le 0,$

and

$(1 + e^{-w})\, k''(w) + (1 - e^{-w})\, k'(w) \ge 0,$

easily obtained by derivation; the corresponding conditions for the general case are, as $\Delta^2_{x,y}\, \Lambda(x,y) \ge 0$:

$k(-\infty) = k(+\infty) = 1,$

$(1 + e^w)\, k(w)$ a non-decreasing function,

$(1+e^{-w})k(w)$ a non-increasing function,

and

$\Delta^2_{x,y}\,[(e^{-x}+e^{-y})\,k(y-x)] \le 0.$

Note that as $(1+e^{w})\,k(w)$ and $(1+e^{-w})\,k(w)$ are monotonic (non-decreasing and non-increasing), we know that k'(w) exists everywhere except for a denumerable set of points of lR.

We will follow the proof of Tiago de Oliveira (1962/63).

As is well known, $\Lambda(x,y)$ is such that

1) $\Lambda(x,+\infty)=\Lambda(x), \Lambda(+\infty,y)=\Lambda(y)$ (margin conditions)

2) $\Delta^2\,\Lambda(x,y)\ge 0.$ (non-negativity condition)

As $\Lambda(x,y)$ is a continuous function, k(w) is also a continuous function. Condition 1) is easily seen to be equivalent to

I) $k(-\infty)=k(+\infty)=1.$

Consider condition 2). It can be written as

$$\exp\{-(e^{-x}+e^{-y})\,k(y-x)\}+\exp\{-(e^{-\xi}+e^{-\eta})\,k(\eta-\xi)\}$$
$$\ge \exp\{-(e^{-x}+e^{-\eta})\,k(\eta-x)\}+\exp\{-(e^{-\xi}+e^{-y})\,k(y-\xi)\}$$
$$\text{for } x\le\xi, y\le\eta.$$

We will show that 2) is equivalent to

II) $(1+e^{w})\,k(w)$ a non-decreasing function,

III) $(1+e^{-w})\,k(w)$ a non-increasing function,

IV) $\{(e^{-x}+e^{-y})\,k(y-x)+(e^{-\xi}+e^{-\eta})\,k(\eta-\xi)\}$
$$\ge\{(e^{-x}+e^{-\eta})\,k(\eta-x)\}+\{(e^{-\xi}+e^{-y})\,k(y-\xi)\}$$
$$\text{for } x\le\xi, y\le\eta.$$

II) and Ill) are obtained taking, respectively, $y\to-\infty$ and $x\to+\infty$. To show IV), let us put

$Z=\exp(-e^{-\xi}), A=(e^{-(x-\xi)}+e^{-(y-\xi)})\,k(y-x),$

$\mathrm{B} = (1 + \mathrm{e}^{-(\eta-\xi)})\, \mathrm{k}(\eta - \xi), \mathrm{C} = (\mathrm{e}^{-(\eta-\xi)} + \mathrm{e}^{-(\eta-\xi)})\, \mathrm{k}(\eta - \mathrm{x}),$

$\mathrm{D} = (1 + \mathrm{e}^{-(y-\xi)}\, \mathrm{k}(y - \xi).$

Conditions 2) and IV) are written as $(\mathrm{A},\mathrm{B},\mathrm{C},\mathrm{D} \geq 0)$

2) $Z^{-A} + Z^{-B} \geq Z^{-C} + Z^{-D}$

IV) $A + B \geq C + D$.

Letting ξ vary with x, y, η but such that $x - \xi\,, y - \xi\,, \eta - \xi$ are fixed, A, B, C, D are also fixed. We then have to prove the full equivalence of 2) and II) under the new form for fixed A, B, C, D and $0 \leq Z \leq 1$.

Let us then consider the function $f(Z) = Z^{-A} + Z^{-B} - Z^{-C} - Z^{-D}$. As $f(Z) \geq 0$ by 2) and $f(1) = 0$ we have $f'(1) \leq 0$ so that IV) is true.

Let us now prove the converse, i.e., that II), III) and IV) imply 2). Conditions II) and III) give $A \geq D, C \geq B$.

Let us prove now 2). Supposing IV) to be true, we have

$Z^{-A}.Z^{-B} \geq Z^{-C}.Z^{-D}$ and thus

$$\max_{\substack{A+B\leq C+D\\ B\leq C, D\leq A}} (Z^{-C} + Z^{-D}) \leq \max_{B\leq C\leq A} \left(Z^{-C} + \frac{Z^{-A}\ Z^{-B}}{Z^{-C}}\right).$$

The RHS function of $C \in [B, A]$ has a minimum at $C = \frac{A+B}{2}$ and the maximum value in the interval is attained at $C = A$ and $C = B$ with the value $Z^{-B} + Z^{-A}$. We have then proved $Z^{-A} + Z^{-B} \geq Z^{-C} + Z^{-D}$ as desired.

If we are dealing with minima with exponential margins, corresponding to the transformation $\xi = e^{-x}, \eta = e^{-y}$, the asymptotic distribution for $\xi\,, \eta \geq 0$ takes the form $\Psi(x, y) = 1 - e^{-\xi} - e^{-\eta} + e^{-(\xi+\eta)A(\eta/(\xi+\eta))}$ or the survival function $S(\xi\,, \eta) = e^{-(\xi+\eta)A(\eta/(\xi+\eta))}$. *It can be shown that the conditions for* $k(.)$ *are equivalent to*

$A(0) = A(1) = 1,$

$0 \leq -A'(0), A'(1) \leq 1$

and $\quad A(u) convex\ in\ [0,1].$

The functions $k(w)$ and $A(u)$ are related by $k(w) = A\left(\frac{1}{1+e^{w}}\right)$ or $A(u) = k(\log\frac{1-u}{u})$. We have $A(u) = 1$ for independence and $A(u) = \max(u, 1-u)$ in the diagonal case, with, also, $\max(u, 1-u)\, A(u) \leq 1$.

Another formulation is

$A(0) = A(1) = 1,$

$max(u, 1-u) \leq A(u) \leq 1$

and $\quad A(u) convex\ in\ [0, 1].$

10.4 Complementary basic results

Some properties can be ascribed to the set of $k(w)$. *The first one is a symmetry property*, i.e., if $k(w)$ is a dependence function, then $k(-w)$ is also a dependence function. The proof is immediate if we consider the conditions in the differentiable case (where a planar density does exist) and slightly longer in the general case. If $k(w) = k(-w)$ then (X, Y) is an *exchangeable pair* and $\Lambda(x, y) = \Lambda(y, x)$.

Also it is immediate that if $k_1(w)$ and $k_2(w)$ are dependence functions, any mixture $\theta\, k_1(w) + (1-\theta)\, k_1(w), 0 \leq \theta \leq 1$ is also a dependence function. *The set of dependence functions is, then, convex*. And this convexity property

$\Lambda_1^{\theta}(x, y) = \Lambda(x, y). \Lambda_2^{1-\theta}(x, y)$

is very useful for obtaining models: the mixed model, as well as the Gumbel model, are such examples. We will call this way of modelling the *mix-technique*.

As a generalization, it can be observed that if $G(u)$ is a distribution function in $[0, 1]$, then $k(w)$ given by $(1 + e^{-w})\, k(w) = \int_0^1 (1 + e^{-w/u})^u\ d\, G(u)$ is also a dependence function.

Another method of generating models is the following *max-technique*. Let (X, Y) be an extreme random pair, with dependence function $k(w)$ and reduced Gumbel margins, and consider the new random pair $(\tilde{X}, \tilde{Y})$ with $\tilde{X} = \max(X + a, Y + b), \tilde{Y} = \max(X + c, Y + d)$. To have reduced Gumbel margins we must have $(e^a + e^b)\,k(a-b) = 1$ and $(e^c + e^d)\,k(c-d) = 1$. Then we have

$$\tilde{k}(w) = \frac{[e^{\max(a+w,c)} + e^{\max(b+w,d)}]\,k[\max(a+w,c) - \max(b+w,d)]}{1+e^w}$$

with (a,b) and (c,d) satisfying the conditions given before. This *max-technique* will be used towards the end of the paper to generate the biextremal and natural models.

Let us stress that independence has a very important position as a limiting situation. If we denote by $P(u,v)$ the function defined by $\text{Prob}\{X > x, Y > y\} = P(F(x,+\infty), F(+\infty,y))$, Sibuya (1960) has shown that the necessary and sufficient condition to have limiting independence is that $P(1-s,1-s)/s \to 0$ as $s \to 0$. He also showed that the necessary and sufficient condition for having the diagonal case as a limiting situation is that $P(1-s,1-s)/s \to 1$ as $s \to 0$. With the first result we can easily show that the maxima of the binormal distribution has independence as a limiting distribution if $|\rho| < 1$.

Also Geffroy (1958/59) showed that a sufficient condition for limiting independence is that

$$\frac{1+F(x,y)-F(x,w_y)-F(w_x,y)}{1-F(x,y)} \to 0 \text{ as } x \to \overline{w}_x \text{ and } y \to \overline{w}_y,$$

$\overline{w}_x$ and $\overline{w}_y$ being the right end-points of the supports of X and Y.

Sibuya conditions (and Geffroy sufficient condition) are easy to interpret: we have limiting independence if $\text{Prob}\{X > x, Y > y\}$ is a vanishing summand of $\text{Prob}\{X > x \text{ or } Y > y\}$ and the diagonal case as limit if $\text{Prob}\,\{X > x, Y > y\}$ is the leading summand of $\text{Prob}\{X > x \text{ or } Y > y\}$.

We have asymptotic independence when $\bar{S}(\xi,\eta) = \xi,\eta$, that is, when $\bar{S}_0^k(\xi^{1/k},\eta^{1/k}) \to \xi,\eta$. The sufficient condition for independence given by Geffroy can be proved in a simpler way, as follows. It is equivalent to:

$$\textit{If } \frac{1-\bar{S}_0(\xi,\eta)}{2-\xi-\eta} \to 1 \qquad \textit{when } \xi,\eta \to 1 \qquad \textit{then } \bar{S}(\xi,\eta) = \xi\,\eta.$$

As $\bar{S}_0(\xi,\eta) = 1-(2-\xi-\eta)\,h(\xi,\eta)$, with $h(\xi,\eta)\to 1$ when $\xi,\eta\to 1$, we have $\bar{S}_0^k\,(\xi^{1/k},\eta^{1/k}) = \{1-(2-\xi^{1/k}-\eta^{1/k})\,h\left(\xi^{1/k},\eta^{1/k}\right)\}^k \to \xi\,\eta$ because

$$h\left(\xi^{1/k},\eta^{1/k}\right)^k \to 1 \text{ when } \eta\to\infty.$$

Its meaning is straightforward, and, in general, we can expect extremal pairs to be asymptotically independent. Mardia (1964) proved that the sample extremes of a bivariate pair are independent under very general conditions.

A simple example now shows that we can have non-independence cases. Let (X,Y,Z) be three independent random variables with the same distribution function F. The distribution function of the pair $(\max(X,Z),\max(Y,Z))$is $P(x,y) = \text{Prob}\{\max(X,Z)\le x,\max(Y,Z)\le y\} = F(x)\,F(y)\,F(\min(x,y))$ and the one for the maxima in a sample of k pairs it is $P_k(x,y) = P^k(x,y)$. Its margins are

$$P_k(x,+\infty) = F^{2k}(x), P_k(+\infty,y) = F^{2k}(y).$$

Supposing now that for λ_k and $\delta_k > 0$, we have

$$F^k(\lambda_k+\delta_k\,z)\to\Lambda(z)$$

we obtain

$$P_k(\lambda_{2k}+\delta_{2k}\,x,\lambda_{2k}+\delta_{2k}\,y)$$

$$= F^k(\lambda_{2k}+\delta_{2k}\,x)\,F^k(\lambda_{2k}+\delta_{2k}\,y)\,F^k(\lambda_{2k}+\delta_{2k}\min(x,y))$$

and we have

$$P_k(\lambda_{2k}+\delta_{2k}\,x,\lambda_{2k}+\delta_{2k}\,y)\to\Lambda^{1/2}(x)\,\Lambda^{1/2}(y)\,\Lambda^{1/2}(\min(x,y)),$$

a case with does not correspond to independence.

It can be remarked that $P(x,y)$, as well as the limit, does not have a planar density.

Sometimes it is usual to study the equidistribution and equidensity curves (if they exist). The latter are very difficult to deal with so that we will restrict ourselves to the former.

The equations of the equidistribution curves

$\Lambda(x, y) = \alpha \qquad (0 < \alpha < l)$

have the form

$(e^{-x} + e^{-y})\, k(y - x) = -\log \alpha \quad (0 < -\log \alpha < +\infty).$

From these equations we see that the translations parallel to the first diagonal

$x \to x + a = x'$

$x \to y + a = y'$

transform the first curve in the other equidistribution curve

$(e^{-x'} + e^{-y'})\, k(y' - x') = (-\log \alpha)\, e^{-a}.$

Consequently we can centre the study of those curves in the median curve $\Lambda(x, y) = 1/2$, that is

$(e^{-x} + e^{-y})\, k(y - x) = \log 2.$

As $k(-\infty) = k(+\infty) = 1$, the asymptotes to the median curves are

$x = -\log\log 2$ and $y = -\log\log 2.$

As these asymptotes are independent of the dependence function, the equidistribution curves behave similarly for large values of x or y, as could be expected, and, then, like the median curve for independence. The relations deduced from Fréchet inequalities yield the double inequality

$e^{-\min(x,y)} \le \log 2 \le e^{-x} + e^{-y},$

showing that any median curve is, necessarily, between the median curves corresponding to the independence

$e^{-x} + e^{-y} = \log 2$

and to the diagonal case

$\min(x, y) = -\log\log 2$

which degenerates in part of the asymptotes.

Consider now the random variable $W = Y - X$, the difference of reduced extremes. We have, evidently, $M(W) = 0$ and $V(W) = \frac{\pi^2}{3}(1 - \rho)$.

Recall that as $V(X) = V(Y) = \pi^2/6$ the covariance $C(X,Y) = \frac{\pi^2}{6}\,\rho$ exists as well as the correlation coefficients ρ.

When $\rho = 1$ we have $V(W) = 0$; W is a random variable almost surely equal to zero and we have $Y = X$ almost surely. We then have the diagonal case.

Let us denote by $D(w) = \text{Prob}\{Y - X \le w\}$; we have

$$D(w) = \int_{-\infty}^{+\infty} d\,x \int_{-\infty}^{x+w} d\,y\, \frac{\partial^2 \Lambda}{\partial\,x\,\partial\,y} = \int_{-\infty}^{+\infty} d\,x\; \Lambda_1(x, x+w)$$

where

$$\Lambda_1(x,y) = \frac{\partial\,\Lambda(x\,,y)}{\partial\,x} = \Lambda(x,y)\{e^{-x}k(y-x) + (e^{-x}+e^{-y})\,k'(y-x)\}$$

so that

$$D(w) = \int_{-\infty}^{+\infty} d\,x\; e^{-e^{-x}(1+e^{-w})k(w)} e^{-x}\,\{k(w) + (1+e^{-w})k'(w)\}$$

$$= \frac{1}{1+e^{-w}} + \frac{k'(w)}{k(w)}$$

as is obvious.

The relations between $D(w)$ and $k(w)$ are immediate. Integration between 0 and w of the differential relation gives

$$(1+e^{w})\,k(w) = 2\,k(0)\exp\int_0^w D(w)\,d\,w,$$

and letting $w \to -\infty (k(-\infty) = 1)$ we obtain

$$k(w) = \frac{\exp\int_{-\infty}^{w} D(w)\,d\,w}{(1+e^{w})}.$$

Note that $\int_{-\infty}^{0} D(w)\,d\,w$ and $\int_0^{+\infty}(1 - D(w))\,d\,w$ exist and are equal because W has mean value zero.

The conditions for $k(w)$ are immediately translated into conditions for $D(w)$ so that

$\mathrm{D}(\mathrm{w}) = \mathrm{Prob}\{\mathrm{Y} - \mathrm{X} \leq \mathrm{w}\}$

verifies the following relations (apart from being a distribution function):

$\int_{-\infty}^{0} \mathrm{D}(\mathrm{w})\, \mathrm{d\,w} = \int_{0}^{+\infty} [1 - \mathrm{D}(\mathrm{w})]\, \mathrm{d\,w}$ or $\int_{-\infty}^{+\infty} \mathrm{d\,w}\, \mathrm{D}(\mathrm{w}) = 0$

and

$\mathrm{D}'(\mathrm{w}) \geq \mathrm{D}(\mathrm{w})\, [1 - \mathrm{D}(\mathrm{w})]$.

We can also write

$\mathrm{D}(\mathrm{w}) = \mathrm{k}'(\mathrm{w})/\mathrm{k}(\mathrm{w}) + \mathrm{L}(\mathrm{w})$,

where $\mathrm{L}(\mathrm{w}) = (1 + \mathrm{e}^{-\mathrm{w}})^{-1}$ is the standard logistic distribution function.

We have $\mathrm{D}(\mathrm{w}) = \mathrm{L}(\mathrm{w})$ in the independence case and $\mathrm{D}(\mathrm{w}) = \mathrm{H}(\mathrm{w})$ where $\mathrm{H}(\mathrm{w})$ is the Heaviside jump function at $0, \mathrm{H}(\mathrm{w}) = 0$ if $\mathrm{w} < 0$ and $\mathrm{H}(\mathrm{w}) = 1$ if $\mathrm{w} \geq 0$ — for the diagonal case. In the case of exchangeability $(\mathrm{k}(\mathrm{w}) = \mathrm{k}(-\mathrm{w}))$ we have $\mathrm{D}(\mathrm{w}) + \mathrm{D}(-\mathrm{w}) = 1$, and W has a symmetric distribution.

As said before, $\mathrm{k}'(\mathrm{w})$ exists everywhere except for a denumerable set of points and thus $\mathrm{D}(\mathrm{w})$ is defined everywhere: directly almost everywhere and at the discontinuity set by the right-continuity of the distribution function $\mathrm{D}(.)$.

Let us denote by $\mathrm{ID}(\mathrm{w}) = \int_{-\infty}^{\mathrm{w}} \mathrm{D}(\mathrm{t})\, \mathrm{d\,t}$; we have seen that

$\mathrm{k}(\mathrm{w}) = \frac{\mathrm{e}^{\mathrm{ID}(\mathrm{w})}}{1+\mathrm{e}^{\mathrm{w}}}$.

The conditions on the distribution function $\mathrm{D}(\mathrm{w})$, in the case of the existence of planar density, are then

$$\int_{-\infty}^{+\infty} \mathrm{w\,d}\;\mathrm{D}(\mathrm{w}) = 0$$

or equivalently

$$\int_{-\infty}^{+\infty} \big(\mathrm{H}(\mathrm{w}) - \mathrm{D}(\mathrm{w})\big)\mathrm{d\,w} = 0$$

and

$D'(w) \geq D(w)(1 - D(w))$.

More generally, the last condition is substituted by

$\Delta^2_{x,y}[\exp(ID(y - x) - y] \leq 0$

which is equivalent to

$e^{ID(w)} - e^{ID(w-\alpha)} \leq e^{-\beta}[e^{ID(w+\beta)} - e^{ID(w+\beta-\alpha)}]$

with $\alpha, \beta \geq 0$, equivalent

$e^{ID(w)}(1 - D(w)) \leq e^{ID(w-\alpha)}(1 - D(w - \alpha))$

or that

$e^{ID(w)}(1 - D(w))$

is a non-increasing function.

Note that $ID(-\infty) = 0, ID(w) - w \to +\infty, ID(w + \beta) - ID(w) \leq \beta (\beta > 0), e^{ID(w)} - e^{ID(w-\alpha)} \leq e^{w} - e^{w-\alpha} (\alpha \geq 0), \max(0, w) \leq ID(w) \leq \log(1 + e^{w})$ (from Boole-Fréchet inequalities), $\int_{-\infty}^{0} D(t)\, d\, t = \int_{0}^{+\infty}(1 - D(t))\, d\, t$ (from the null mean value), $D(w)\, e^{ID(w)} \leq e^{w}$ and $(1 - D(w))\, e^{ID(w)} \leq 1$ (by convenient integration of $D' \geq D(1 - D)$).

Let $\underline{w}$ and $\overline{w}$ ($\underline{w} \leq 0 \leq \overline{w}, \overline{w} = \underline{w} = 0$ in the diagonal case and $\underline{w} < 0 < \overline{w}$ in the other cases) be the left and right end-points of $D(w)$.

For $w > \overline{w}$, if $\overline{w} < \infty$, we have

$ID(w) = \int_{-\infty}^{w} D(t)\, d\, t = \int_{+\infty}^{\overline{w}} D(t)\, d\, t + w - \overline{w}$;

as

$$ID(w) - w = \int_{-\infty}^{\overline{w}} D(t)\, d\, t - \overline{w} \to 0$$

we have

$\overline{w} = \int_{-\infty}^{\overline{w}} D(t)\, d\, t,$

a result that is also true $\overline{w} = +\infty$.

Consider now, for $\underline{w} > -\infty$,

$$\int_{\underline{w}}^{+\infty}(1 - D(t))\,d\,t = \int_{\underline{w}}^{0}(1 - D(t))\,d\,t + \int_{0}^{+\infty}(1 - D(t))d\,t =$$

$$-\underline{w} - \int_{\underline{w}}^{0} D(t)\,d\,t + \int_{0}^{+\infty}(1 - D(t))d\,t = -\underline{w}$$

which is also true $\underline{w} = -\infty$.

If both $\underline{w}$ and $\overline{w}$ are finite we have

$$\overline{w} = \int_{\underline{w}}^{w} D(t)\,d\,t \quad \text{and} \quad \underline{w} = -\int_{\underline{w}}^{\overline{w}}(1 - D(t))\,d\,t.$$

The symmetry condition (exchangeability) $k(w) = k(-w)$ is equivalent to $D(w) + D(-w) = 1$ or $ID(w) = w + ID(-w)$.

From the conditions on $D(w)$ we can give new methods of generating bivariate models of maxima, as follows, essentially for absolutely continuous distribution functions. The most natural is the one that follows.

In $[\underline{w}, \overline{w}]$, the support of $D(.)$, we can define the function $\Psi(w) = \frac{D'(w)}{D(w)\,(1-D(w))} \geq 1$. Notice that as a consequence every point of $[\underline{w}, \overline{w}]$ is a point of increase of $D(.)$, and thus the quantiles are uniquely defined. From the absolute continuity we have $D(\underline{w}) = 0$ and $D(\overline{w}) = 1$. Note that $\underline{w} < 0 < \overline{w}$, as the mean value is zero.

Let us fix v as the median. Integrating the differential equation $D' = D(1 - D)\Psi$ we get, with $I\Psi(w) = \int_{-\infty}^{w} \Psi(t)\,d\,t$,

$$D(w|v) = \frac{1}{1+e^{-(I\Psi(w)-I\Psi(v))}},$$

and thus the relation between $D(w|v)$ and $D(w|v')$ is

$$D(w|v') = \frac{D(w|v)}{D(w|v) + e^{-K}(1 - D(w|v))}$$

with $K = \int_{v'}^{v} \Psi(t)\,d\,t$ and the same support.

As $\Psi(w) \geq 1$ it is immediate that $\int_{v}^{w} \Psi(t)\, d\, t \geq (w - v)$ for $w > v$. We get, thus

$$D(w|v) \geq \frac{1}{1 + e^{-(w-v)}}$$

and for $w < v$, analogously, we obtain

$$D(w|v) \leq \frac{1}{1+e^{-(w-v)}}\,,$$

results not known previously.

Having fixed, temporarily, the median at $v(\underline{w} < v < \overline{w})$, the condition

$\int_{-\infty}^{+\infty} w\, d\, D(w|v) = \int_{-\infty}^{+\infty} [H(w) - D(w|v)]\, d\, w = 0$ reads as

$I(K) = 0$ where

$$I(K) = \int_{\underline{w}}^{\overline{w}} [H(w) - \frac{D(w|v)}{D(w|v)+e^{-K}(1-D(w|v))}]\, d\, w.$$

As $I(K) \to \underline{w}$ when $K \to +\infty, I(K) \to \overline{w}$ when $K \to -\infty$ and $I(K)$ is increasing, we see that the solution of $I(K) = 0$ is unique. The unique solution K_0 of $I(K) = 0$ leads to

$$\int_{v'}^{v} \Psi(t)\, d\, t = K_0$$

and, as $\Psi(w) \geq 1$, the only solution for $D(.)$ is immediately obtained. It is thus sufficient to give an integrable function $\Psi(w) \geq 1$ in every finite interval $[\underline{w}\,, \overline{w}]$ to obtain the unique $D(.)$ related to it. For simplicity we can take $v = 0$.

10.5 Correlation results

The results concerning correlation that follow and the regression results to be given later, to a certain extent, illuminate some other features of the situation.

Let us now compute some widely used correlation coefficients.

The classical correlation coefficient always exists as shown before and is $\rho = \frac{6}{\pi^2}\, C(X, Y)$.

A known expression for $C(X, Y)$ is

$$C(X,Y) = \int_{-\infty}^{+\infty}\int [F(x,y) - F(x,+\infty)\, F(+\infty,y)]\, d\,x\ \, d\,y$$

which, in our case, takes the form

$$C(X,Y) = \int_{-\infty}^{+\infty}\int [\Lambda(x,y) - \Lambda(x)\,\Lambda(y)]\, d\,x\ \, d\,y,$$

and because of the positive association we have $C(X,Y) \geq 0$, substituting the expressions of $\Lambda(x,y), \Lambda(x)$ and $\Lambda(y)$ and using the change of variables

$$e^{-v} = e^{-x} + e^{-y}$$

$$w = y - x,$$

we obtain

$$C(X,Y) = \int_{-\infty}^{+\infty} d\,w \int_{-\infty}^{+\infty} d\,v\, [e^{-e^{-v}k(w)} - e^{-e^{-v}}].$$

Let us now compute the inner integral. It can be written as

$$\int_0^{+\infty}(1 - e^{-e^{-v}})d\,v - \int_{-\infty}^{0} e^{-e^{-v}}\, d\,v + \int_{-\infty}^{0} e^{-e^{-v}k(w)}d\,v - \int_0^{+\infty}(1 - e^{-e^{-v}k(w)})\, d\,v.$$

The first two integrals, being an expression for $M(X)$, add to γ. Changing v to $v = z + \log k(w)$ in the two last integrals, we have after simple computations for the inner integral

$$\gamma + \int_{-\infty}^{0} e^{-e^{-z}}\, d\,z - \int_0^{+\infty} (1 - e^{-e^{-z}})\, d\,z - \int_{-\log k(w)}^{0} d\,z = -\log k(w)$$

because the two first integrals add to $-\gamma$.

We then have the expression:

$$C(X,Y) = -\int_{-\infty}^{+\infty} \log k(w)\, d\,w \geq 0$$

and thus the classical correlation coefficient for $\Lambda(x,y)$ is

$$\rho = -\frac{6}{\pi^2}\int_{-\infty}^{+\infty} \log k(w)\, d\,w\,;$$

when needed we will write $\rho(k)$ to indicate the dependence function.

From the expression of $D(w)$ we can obtain, once more in the differentiable case, also the expression $\rho = -\frac{6}{\pi^2}\int_{-\infty}^{+\infty} \log k(w)\, d\, w$. Let us give the supplementary proof.

Let us recall first that the existence of the variance $V(W)$ implies that $\lim_{w\to-\infty} w^2\, D(w) = 0$ and $\lim_{w\to\infty} w^2\,(1 - D(w)) = 0$, and, as the same happens for the logistic $1/(1 + e^{-w})$, those conditions are equivalent to $\lim_{w\to+\infty} w^2 \frac{k'(w)}{k(w)} = 0$. As the variance of the logistic ($\rho = 0$) is $\pi^2/3$, we have the general relation ($M(W)$ being zero)

$$V(W) = \frac{\pi^2}{3}(1-\rho) = \int_{-\infty}^{+\infty} w^2\, d\, D(w) = \int_{-\infty}^{+\infty} w^2\, d(\frac{1}{1+e^{-w}}) + \int_{-\infty}^{+\infty} w^2\, d(\frac{k'(w)}{k(w)})$$

so that

$$-\frac{\pi^2}{3}\rho = \int_{-\infty}^{+\infty} w^2\, d(\frac{k'(w)}{k(w)}).$$

Integration by parts gives

$$\frac{\pi^2}{3}\rho = 2\int_{-\infty}^{+\infty} w\ \frac{k'(w)}{k(w)},$$

the integrated parts being zero as a consequence of

$$\lim_{w\to\pm\infty} w^2 \frac{k'(w)}{k(w)} = 0.$$

The last integral is equal, by integration by parts, to

$$\int_{-\infty}^{+\infty} \log k\,(w)\, d\, w,$$

with the integrated part $w \log k(w) \to 0$ when $w \to \pm\infty$, as follows from the Fréchet derived inequalities.

We can summarize the results by giving the equivalent expressions for ρ:

$$\rho = 1 - \frac{3}{\pi^2}\int_{-\infty}^{+\infty} w^2\, d\, D(w) = \frac{6}{\pi^2}\int_{-\infty}^{+\infty} w\,\{D(w) - L(w)\}\, d\, w$$

$$= \frac{6}{\pi^2}\,.\int_{-\infty}^{+\infty}(IL(t) - ID(t))\, d\, t = -\frac{6}{\pi^2}\int_{-\infty}^{+\infty} \log k(w)\, d\, w.$$

As $k(w) \leq 1 (-\log k(w) \geq 0)$ we have $0 \leq \rho$, as could be expected from the positive association. It is very easy to show that for the diagonal case

$$k_D(w) = \frac{\max(1, e^w)}{1 + e^w}$$

we have $\rho = 1$. The value of ρ does not identify the dependence function (or the distribution): ρ is the same for $k(w)$ and $k(-w)$. But $\rho = 0$, as $k(w) \leq 1$, implies $k(w) = 1$, or independence. Now writing ρ under the form

$$\rho = 1 - \frac{6}{\pi^2} \int_{-\infty}^{+\infty} \log \frac{k(w)}{k_D(w)} \, d\, w,$$

we see, analogously, that $\rho = 1$ as $k(w) \geq k_D(w)$ implies $k(w) = k_D(w)$, or the diagonal case.

For the non-parametric correlation coefficients used, see Fraser (1957) and Konijn (1950) for details.

With the expression

$$\chi = 12 \int_{-\infty}^{+\infty} \int \Lambda(x, y) \, d\, \Lambda(x) \, d\, \Lambda(y) - 3$$

for the grade correlation coefficient we immediately find that the grade correlation coefficient has the expression

$$\chi = 12 \int_{-\infty}^{+\infty} \frac{e^w}{(1+e^w)^2 \, (1+k(w))^2} \, d\, w - 3.$$

It is immediate also that as $k(w) \leq 1$ the integral in the final expression for χ is greater than

$$\int_{-\infty}^{+\infty} \frac{e^w}{(1+e^w)^2} \cdot \frac{d\, w}{4} = \frac{1}{4} \text{ so that } \qquad \chi \geq 0.$$

Other expressions for χ are

$$\chi = 12 \int_{-\infty}^{+\infty} \frac{e^w}{(1+e^w)^2 \, (1+k(w))^2} \, d\, w - 3 = 12 \int_{-\infty}^{+\infty} \frac{e^w}{(1+e^w+e^{ID(w)})^2} \, d\, w - 3.$$

For the difference-sign correlation coefficient

$$\tau = 4 \int_{-\infty}^{+\infty} \int \Lambda(x, y) \, d\, \Lambda(x, y) - 1,$$

in the case of differentiability, we show that the expression for the difference-sign correlation is

$$\tau = \int_{-\infty}^{+\infty} \frac{k'(w)^2}{k(w)} d\, w - 2 \int_{-\infty}^{+\infty} \frac{e^w}{(1+e^w)^2} \log k(w) d\, w,$$

which can also take the form

$$\tau = 1 - \int_{-\infty}^{+\infty} D(w)(1 - D(w))\, d\, w = \int_{-\infty}^{+\infty} (D^2(w) - L^2(w)) d\, w.$$

By integration by parts of $\int_{-\infty}^{+\infty} D(w)\big(1 - D(w)\big)\, d\, w$ as $\int_{-\infty}^{+\infty} w\, d\, D(w) = 0$, we get the simple expression

$$\tau = 1 - \int_{-\infty}^{+\infty} w\, d\;\, D^2(w).$$

Finally for the medial correlation coefficient

$$v = 4[\Lambda(\tilde{\mu}, \tilde{\mu}) - \Lambda(\tilde{\mu})\, \Lambda(\tilde{\mu})] = 4\, \Lambda(\tilde{\mu}, \tilde{\mu}) - 1,$$

$\tilde{\mu} = -\log\log 2$ being the median, we see that the medial correlation coefficient has the expression $v = 4^{1-k(0)} - 1 (\geq 0)$.

The probability of concordance, that is, the probability that the two pairs (X_1, Y_1) and (X_2, Y_2) are such that $X_1 - X_2$ and $Y_1 - Y_2$ have the same sign, being given by $(1 + \tau)/2$, is always greater than $1/2$.

As an index of dependence we can, also, use

$$\Delta(k) = \sup_{x,y} |\, \Lambda(x, y) - \Lambda(x)\, \Lambda(y)| = \sup_{x,y} [\Lambda(x, y) - \Lambda(x)\Lambda(y)]$$

as follows from positive association.

Computations show easily that if $k(w)$ is differentiable and w_0 is such that $k'(w_0) = 0$ the value of $\Delta(k)$ is, with $k_0 = k(w_0)$,

$$\Delta(k) = k_0^{k_0/(1-k_0)} - k_0^{1/(1-k_0)};$$

for $k(w) = k(-w)$ we have $w_0 = 0$. It is very easy to see from the expression of $k'(w)$ has only an interval of solutions so that $k(w)$ has the same value for all w_0 in the interval (a minimum). The same result an be obtained directly without recourse to the hypothesis of differentiability, as follows.

Denoting by $Z = \Lambda(x)\,\Lambda(y)$, the expression of $\Delta(k)$ can be written (as $k(w) \leq 1$)

$$\sup_{\substack{-\infty \leq w \leq +\infty \\ 0 \leq Z \leq 1}} (Z^{k(w)} - Z)$$

which for fixed $Z(\leq 1)$ has the maximum for the minimum $k_0 = k(w_0)$ of $k(w)$. The maximum of $Z^{k(w_0)} - Z$ is given by the previous value. Then the index of dependence

$$\Delta(k) = \sup_{x,y} [\Lambda(x,y) - \Lambda(x)\,\Lambda(y)]$$

has the value $\Lambda(k) = k_0^{k_0/(1-k_0)} - k_0^{1/(1-k_0)}$, k_0 being the minimum of k.

It can be recalled that the classical correlation coefficient is linearly invariant (independent of the margin parameters) and that all the others are transformation invariant (non-parametric).

Consider, finally, the case for bivariate minima with standard exponential margins. Then the mean values and variances are equal to 1; the covariance (evidently equal to the correlation coefficient) is, using the same formula as the one at the beginning, transformed to survival functions, $(\xi, \eta \geq 0)$

$$\rho = C(\xi, \eta) = \int_0^{+\infty} \int (S(\xi, \eta) - S_0(\xi, \eta)\, d\,\xi\ d\,\eta$$

where $S_0(\xi, \eta) = e^{-\xi}\, e^{-\eta}$ is the survival function for independence. By the change of variables $s = \xi + \eta$, $u = \eta/(\xi + \eta)$ we get

$$\rho = \int_0^1 d\,u \int_0^{+\infty} (e^{-sA(u)} - e^{-s})\, s\, d\,s = \int_0^1 (1/A^2(u) - 1) d\,u$$

$$= \int_0^1 A^{-2}(u)\, d\,u - 1 \geq 0 \ \text{ as } \ (1/2 \leq) \max(u, 1-u) \leq A(u) \leq 1$$

with the value $\rho = 0$ for independence $(A(u) = 1)$ and $\rho = 1$ for the diagonal case $(A(u) = \max(u, 1-u))$ as should be expected. Recall that ρ is linearly invariant although in this case only scale invariance matters.

The other dependence indicators, being non-parametric, have the same expression, after transformation from $k(w)$ to $A(u)$.

We have then:

for the grade correlation $\chi = 12 \int_0^1 \frac{d\,u}{(1+A(u))^2} - 3$

for the difference-sign correlation

$\tau = \int_0^1 (\frac{A'(u)}{A(u)})^2\, u(1-u)\, d\,u - 2 \int_0^1 \log A\,(u)\, d\,u$;

for the medial correlation $v = 4^{1-A(1/2)} - 1$;

for the index of dependence $\Lambda(A) = A_0^{A_0/(1-A_0)} - A_0^{1/(1-A_0)}$ where $A_0(\geq 1/2)$ denotes the minimum of $A(u)$.

As $\xi = e^{-X}, \eta = e^{-Y}, u = (1+e^{w})^{-1}$, we see that $B(u) = \text{Prob}\{\frac{\eta}{\xi+\eta} \leq u\} = \text{Prob}\{Y - X \geq \log\frac{1-u}{u}\} = 1 - D(\log\frac{1-u}{u}) = u + \frac{u(1-u)A'(u)}{A(u)}$ and $\tau = 1 - \int_0^1 \frac{B(u)(1-B(u))}{u(1-u)} d\,u = 1 - \int_0^1 \log\frac{1-u}{u} d\,B^2(u)$ because $\int_0^1 \log\frac{1-u}{u} d\,B(u) = 0$ as follows from $\int_{-\infty}^{+\infty} w\, d\,D(w) = 0$.

10.6 Regression results

Let us now discuss some results on regression in the case of reduced Gumbel margins. For convenience, we will only consider the regression of Y on X, the regression of X on Y being dealt with in the same way with the substitution of $k(w)$ by $k(-w)$ as said before. The linear regression line for reduced margins is evidently $L_y(x) = \gamma + \rho(k)\,(x-\gamma), \rho(k)$ being given before.

As $\text{Prob}\{Y \leq y | X = x\} = \Lambda(y|x)$ is given, in the case of existence of a density, by

$$\Lambda(y|x) = \frac{1}{\Lambda'(x)} \frac{\partial\,\Lambda(x,y)}{\partial\,x} = \exp\{e^{-x} - (e^{-x} + e^{-y})\, k(y-x)\} \times \{k(y-x)$$

$$+\left(1 + e^{-(y-x)}\right)k'(y-x)\}$$

the regression line is

$$\bar{y}(x|k) = \int_{-\infty}^{+\infty} y\, d\;\Lambda(y|x)$$

and as $\gamma = \int_{-\infty}^{+\infty} y\, d\;\Lambda(y)$ we get, by integration by parts,

$$\bar{y}(x|k) = \gamma + \int_{-\infty}^{+\infty} [\Lambda(y) - \Lambda(y|x)]\, d\, y$$

$= \gamma + \int_{-\infty}^{+\infty} \{\exp(-e^{-x}e^{-w}) - [k(w) + (1 + e^{-w})k'(w)].$

$\exp\{e^{-x}[(1 + e^{-w})k(w) - 1]\}\, d\, w$

$= \gamma + e^{x} + \int_{-\infty}^{+\infty} [e^{-e^{-x}e^{-w}} - (1 + e^{-w})\, k(w)\, e^{-e^{-x}[(1+e^{-w})k(w)-1]} d\, w.$

The correlation ratio is given by

$$R^2(y|x; k) = \frac{6}{\pi^2} \int_{-\infty}^{+\infty} (\bar{y}(x|k) - \gamma)^2 \, d\, \Lambda(x)$$

$$= \frac{6}{\pi^2} \int_{-\infty}^{+\infty}\int d\, v\, d\, w \left[\frac{e^{-v}\, e^{-w}}{1 + e^{-v} + e^{-w}} - 2 \frac{e^{-v}\big((1 + e^{-w})\, k(w) - 1\big)}{e^{-v} + (1 + e^{-w})\, k(w)} + \right.$$

$$\left. + \frac{\big((1+e^{-v})\, k(v) - 1\big)\big((1+e^{-w})\, k(w) - 1\big)}{(1+e^{-w})\, k(v) + (1+e^{-w})\, k(w) - 1} \right].$$

Median regression can be defined as the solution of the equation $\Lambda(\tilde{y}|x) = 1/2$; it takes the form $\tilde{y}(x|k) = x + \varphi(x)$ where $\varphi(x)$ is given by $e^{-x}(1 + e^{-\varphi})\, k(\varphi) - \log\,[k(\varphi) + (1 + e^{-\varphi})k'(\varphi)] = (\log 2)\, e^{-x}$. In two of the models (logistic and mixed) described below the curves are approximately linear — see Gumbel and Mustafi (1968).

Evidently where we have margin parameters we must substitute x and y by $(x - \lambda_x)/\delta_x$ and $(y - \lambda_y)/\delta_y$ in all the three cases above. For instance, for linear regression we have $\frac{L_y(x) - \lambda_y}{\delta_y} = \gamma + \rho(\frac{x - \lambda_x}{\delta_x} - \gamma)$ or $L_y(x) = \lambda_y + \gamma\, \delta_y + \rho\, \delta_y \frac{x - \lambda_x - \gamma\, \delta_x}{\delta_x}$, which can take the usual form $L_y(x) = \mu_y + \rho \frac{\sigma_y}{\sigma_x}(x - \mu_x)$.

Let us return, from now on, to the study of regression for reduced values of (X, Y).

As is well known, see Cramér (1946), the mean-square error of the linear regression for reduced Gumbel margins is

$$MSE(L_y(X)) = M_{X,Y}(Y - \gamma - \rho(X - \gamma))^2 = \frac{\pi^2}{6}(1 - \rho^2)$$

when the mean-square error of the general regression is

$$MSE(\bar{y}(X)) = M_{X,Y}(Y - \bar{y}(X))^2 = \frac{\pi^2}{6} - M_X(\bar{y}(X) - \gamma)^2 = \frac{\pi^2}{6}(1 - R^2)$$

which is, as known, equal to the variance $V(Y - \bar{Y})$.

The mean-square error reduction in percentage is

$$NL = \frac{MSE(L_y) - MSE(\bar{y})}{MSE(L_y)} = \frac{R^2 - \rho^2}{1 - \rho^2};$$

it can be considered the index of non-linearity, evaluating the improvement in using non-linear regression — see Tiago de Oliveira (1974).

It can observe that in all the cases studied so far, NL is very small and so the improvement is negligible. This is easily understandable because $L_y(x)$ and $\bar{y}(x)$ are, in fact, very different for $x < -2$ or $x > 5$ whose total probability is very low (about .007 or less).

The bounds for k(w) can give bounds for $\bar{y}(x)$, using the expression of $\bar{y}(x|k) - \gamma$ given before, but the bounds are so large that they are not useful. In fact we have

$$-\gamma - 1 + x - e^{e^{-x}} \int_{e^{-x}}^{+\infty} \frac{e^{-t}}{t} d\,t \le \bar{y}(x|k) - \gamma \le e^x(1 - e^{e^{-x}}) + \int_{e^{-x}}^{+\infty} \frac{e^{-t}}{t} d\,t.$$

Let us see how we can, if needed, improve the regression lines, through the use of orthogonal polynomials with respect to $\Lambda(x)$, that we are going to construct.

Let $\{\Psi_p(x)\}$ denote the complete set of orthogonal polynomials with respect to $\Lambda(x)$, i.e.

$$\int_{-\infty}^{+\infty} \Psi_p(x)\,\Psi_q(x)\, d\,\Lambda(x) = \delta_{pq}$$

where δ_{pq} is the Kroneker symbol $(\delta_{pp} = 1, \delta_{pq} = 0 \text{ if } p \neq q)$, $\Psi_p(x)$ having the degree p. Recall that $\{(x - \gamma)^p\}$ is a complete set of polynomials (not orthogonal) in any bounded interval.

Let us take $\Psi_0(x) = 1$ and $\Psi_{k+1}(x) = \sum_0^k a_{kj}\,\Psi_j(x) + b_{k+1}(x - \gamma)^{k+1}$. The orthonormality conditions give for $t = 0,1,\ldots,k+1$

$$\int_{-\infty}^{+\infty} \Psi_{k+1}(x)\,\Psi_t(x)\,d\,\Lambda(x) = \delta_{k+1,t}$$

which for $t = 0, 1, \dots, k$ give

$$0 = \sum_0^k a_{kj}\ \delta_{jt} + b_{k+1} \int_{-\infty}^{+\infty} (x-\gamma)^{k+1}\ \Psi_t(x)\,d\,\Lambda(x) = a_{kt} + b_{k+1}\ \varphi_{k+1,t}$$

and thus,

$\Psi_{k+1}(x) = b_{k+1}(-\sum_0^k \varphi_{k+1,t}\ \Psi_t(x) + (x-\gamma)^{k+1})$.

The normalization equation, for $k+1$, gives

$$\int_{-\infty}^{+\infty} \Psi_{k+1}^2(x)\,d\,\Lambda(x) = 1$$

or

$$b_{k+1}^2 \int_{-\infty}^{+\infty} ((x-\gamma)^{k+1} - \sum_0^k \varphi_{k+1,t}\ \Psi_t(x))^2\,d\,\Lambda(x) = 1$$

or

$b_k^2 = 1/(\mu_{2k+2} - \sum_0^k \varphi_{k+1,t}^2)$;

for simplicity we will take $b_k = (\mu_{2k+2} - \sum_0^k \varphi_{k+1,t}^2)^{-1/2}$ (the positive root for the radical).

We have, immediately,

$\Psi_0(x) = 1, \Psi_1(x) = \sqrt{\frac{6}{\pi^2}}\ (x-\gamma)$,

$$\Psi_2(x) = \frac{6(x-\gamma)^2/\pi^2 - \beta_1\sqrt{6/\pi^2}\ (x-\gamma) - 1}{\sqrt{\beta_1 - \beta_1^2 - 1}}, \text{etc.}$$

where $\beta_1 = \mu_3/\mu_2^{3/2} = 1.13958$ and $\beta_2 = \mu_4/\mu_2^2 = 5.4$, as known.

Suppose that we want to develop $\bar{y}(x) - \gamma = \sum_0^\infty \rho_p\,\Psi_p(x)$.

We have

$$\frac{6}{\pi^2}\, R^2 = M(\bar{y}(x)-\gamma)^2 = \int_{-\infty}^{+\infty} \Big(\sum_0^k \rho_p\ \Psi_p(x)\Big)^2\, d\,\Lambda(x) =$$

$$\sum_{p,q} \rho_p\ \rho_q \int_{-\infty}^{+\infty} \Psi_p(x)\ \Psi_q(x)\, d\,\Lambda(x) = \sum_0^\infty \rho_p^2.$$

Let us now compute the ρ_j. We have

$$\int_{-\infty}^{+\infty} (\bar{y}(x)-\gamma)\,\Psi_j(x)\, d\,\Lambda(x) = \sum_0^\infty \rho_p \int_{-\infty}^{+\infty} \Psi_p(x)\,\Psi_j(x)\, d\,\Lambda(x)$$

$$= \sum_0^\infty \rho_p\, \delta_{pj} = \rho_j, \text{i.e.,}$$

$$\rho_j = \int_{-\infty}^{+\infty} (\bar{y}(x)-\gamma)\,\Psi_j(x)\, d\,\Lambda(x)$$

Thus we have:

$$\rho_0 = \int_{-\infty}^{+\infty} 1.\,(\bar{y}(x)-\gamma)\, d\,\Lambda(x) = 0$$

$$\rho_1 = \int_{-\infty}^{+\infty} \Psi_1(x)\,(\bar{y}(x)-\gamma)\ d\,\Lambda(x) = a_{11} \int_{-\infty}^{+\infty} (x-y)(\bar{y}(x)-\gamma)\, d\,\Lambda(x) =$$

$$= a_{11}\, \frac{\pi^2}{6}\rho = \sqrt{\frac{\pi^2}{6}}.\rho$$

$$\frac{\pi^2}{6} R^2 = \frac{\pi^2}{6}\rho^2 + \sum_2^\infty \rho_j^2 \quad \text{or} \quad R^2 = \rho^2 + \frac{6}{\pi^2} \sum_2^\infty \rho_j^2.$$

For the independence case we get $\bar{y}(x)-\gamma$, and so as $\rho = 0$,

$$\frac{\pi^2}{6} R^2 = \sum_2^\infty \rho_j^2\,, \rho_j = \int_{-\infty}^{+\infty} (\gamma-\gamma)\,\Psi_j(x)\, d\,\Lambda(x) \Rightarrow R^2 = \rho^2 = 0,$$

and for the diagonal case we obtain $\bar{y}(x) = x$, so

$$\rho_j = \int_{-\infty}^{+\infty} (x-\gamma)\,\Psi_j(x)\, d\,\Lambda(x) = \frac{1}{a_{11}} \int_{-\infty}^{+\infty} \Psi_1(x)\,\Psi_j(x)\, d\, x =$$

$$= \frac{1}{a_{11}}\,\delta_{pj} \Rightarrow \rho_j = 0, \text{ except } \quad \rho_1 = \frac{1}{a_{11}} = \sqrt{\frac{\pi^2}{6}},$$

and so

$$\frac{\pi^2}{6} R^2 = \frac{\pi^2}{6} \Rightarrow R^2 = 1.$$

10.7 Miscellaneous results

The results contained in this section were, in part, given in Tiago de Oliveira (1962/63) and (1964). They refer to bivariate pairs with reduced Gumbel margins.

Let us prove that

If (X, Y) is an extremal pair, the distribution of $(Y|X)$ is extremal if and only if (X, Y) is an independent pair.

It is evident that if (X, Y) is an independent pair $(Y|X) = Y$ is an extremal variate. Let us prove the converse.

Denoting by (as in the previous consideration of regression)

$$\Lambda(y|x) = e^{-(e^{-x}+e^{-y})k(y-x)+e^{-x}}\{k(y-x) + (1+e^{-y+x})k'(y-x)\}$$

the distribution function of $(Y|X)$, we must have

$$\Lambda(y|x) = \Lambda(\frac{y-v(x)}{\tau(x)}).$$

This relation takes the form

$$(e^{-x}+e^{-y})\,k(y-x) - e^{-x} - \log\{\,k(y-x) + (1+e^{-(y-x)})\,k'(y-x)\} = e^{-(y-v(x)/\tau(x))}.$$

Taking now $x_1 < x_2$ such that $\tau(x_1) > 0, \tau(x_2) > 0$, putting $y = w + x_1$ and $y = w + x_2$ and subtracting, we obtain

$$(1+e^{-w})\,k(w) = 1 + \alpha\, e^{-\beta w} - \alpha'\, e^{-\beta' w},$$

$\beta = \tau(x_1)^{-1} > 0, \beta' = \tau(x_2)^{-1} > 0$ and $\alpha \geq 0$ and $\alpha' \geq 0$ are functions of x_1 and x_2.

Take $\beta \geq \beta'$; $k(w)$ can be written as

$$k(w) = \frac{1 + e^{-\beta w}(\alpha - \alpha' \, e^{(\beta-\beta')w})}{1 + e^{-w}}$$

and as $k(-\infty) = 1$ we obtain

$$k(-\infty) = \lim_{w \to -\infty} \alpha \; e^{(1-\beta)w} = 1$$

so that $\beta = 1, \alpha = 1$. Consequently we have

$$k(w) = 1 - \alpha' \frac{e^{-\beta' w}}{1+e^{-w}}, \qquad (\beta' \leq 1),$$

Condition II) gives $\alpha' \leq 0$ so that and $\alpha' = 0$ and $k(w) = 1$. The case $\beta' \leq \beta$ is dealt with in the same way.

From the expression $\rho = -\frac{6}{\pi^2} \int_{-\infty}^{+\infty} \log k(w) \, d\, w$, we see that $\rho = 0$ if and only if $k(w) = 1$, that is, if the extremal pair is independent.

An extremal pair (X, Y) *is independent if and only if any pair* $(X_\theta, Y_\theta), X_\theta = \cos\theta X + \mathrm{sen}\,\theta Y + a, Y_\theta = -\mathrm{sen}\,\theta X + \cos\theta Y + b$ *for any* $\theta \neq 0, \pi/2, \pi, 3\pi/2$ *have equal variances.*

The result follows immediately from the fact that

$$V(X_\theta) = \frac{\pi^2}{6}(1 + \rho \,\mathrm{sen}\, 2\,\theta)$$

$$V(Y_\theta) = \frac{\pi^2}{6}(1 - \rho \,\mathrm{sen}\, 2\,\theta).$$

A side result is that

$Z = a + a' X + a'' Y$ *has an extremal distribution when* (X, Y) *is an extremal independent pair if and only if* $a' = 0$ *or* $a'' = 0$.

One half of the proof is straightforward. Let us prove the converse. We can suppose that X, Y and Z have reduced extremal distributions, only having to transform linearly Z if this is not the case (the coefficients of X and Y being proportional to a' and a"). As $M(X) = M(Y) = M(Z) = \gamma$ we obtain

$$Z - \gamma = a'(X - \gamma) + a''(Y - \gamma).$$

The computation of the second and fourth central moments immediately gives

$a'^2 + a''^2 = 1$

$\beta_2(a'^4 + a''^4) + 6a'^2\, a''^2 = \beta_2,$

and so

$(6 - 2\beta_2)a'^2\, a''^2 = 0;$ as $\beta_2 = 5.4$ we have $a' = 0$ or $a'' = 0$.

As an immediate consequence the pair (X, Y) $(X = a + a'\, X_0 + a''\, Y_0, Y = b + b'\, X_0 + b''\, Y_0)$ is an extremal pair, (X_0, Y_0) being an independent extremal pair, if and only if $a'' = b' = 0$ or $a' = b'' = 0$.

Another result, corresponding to the characterization of $\Lambda(x, y)$, *is that* $Z = max(\alpha X + a, \beta Y + b)$ $(\alpha, \beta > 0)$ *is extremal,* (X, Y) *being a reduced extremal pair, if and only if* $\alpha = \beta$.

As $\Lambda(z) = \text{Prob}\{\max(\alpha\, X + a, \beta\, Y + b) \le z\} = \Lambda(\frac{z-a}{\alpha}, \frac{z-b}{\beta})$ we must seek conditions such that

$$\Lambda(\frac{z-a}{\alpha}, \frac{z-b}{\beta}) = \Lambda(\frac{z-\lambda}{\delta})$$

or such that $(z = \lambda + \delta\, v)$

$$\Lambda(v) = \Lambda(\frac{\lambda-a}{\alpha} + \frac{\delta}{\alpha}v, \frac{\lambda-b}{\beta} + \frac{\delta}{\beta}v) = \Lambda(a' + \alpha'\, v, b' + \beta' v),$$

that is

$$e^{-v} = (e^{-a'-\alpha' v} + e^{-b'-\beta' v})\, k(b' + \beta'\, v - a' - \alpha'\, v).$$

Multiplying this equation by e^v and letting $v \to \pm\infty$ we have a limit if and only if $\alpha' = \beta' = 1$ or $\alpha = \beta = \delta$ and then

$$e^{\lambda/\delta} = (e^{a/\delta} + e^{b/\delta})\, k(\frac{a-b}{\delta}),$$

which relates λ and δ.

$\Lambda(x, y|\theta) = exp(-(e^{-x} + e^{-y})\, k(y - x|\theta))$ *cannot have sufficient statistics (of rank 1) if for some* $\theta = \theta_0$ *we have independence.*

As is well known, if $\Lambda(x,y|\theta)$ does have a sufficient statistic for θ (of rank 1), its density

$$\frac{\partial^2 \Lambda(x,y|\theta)}{\partial x \ \partial y} = \exp\{-(e^{-x}+e^{-y})k(y-x|\theta)\}.\{e^{-x}\,P(y-x|\theta)+e^{-2x}\,Q(y-x|\theta)\}$$

with

$$P(w|\theta) = (1+e^{-w})k''(w|\theta) + (1-e^{-w})k'(w|\theta) \geq 0$$

and

$$Q(w|\theta) = \{k(w|\theta)+(1+e^{-w})k'(w|\theta)\}.\{e^{-w}k(w|\theta)-(1+e^{-w})\,k'(w|\theta)\} \geq 0$$

would be of the form

$$\frac{\partial^2 \Lambda(x,y|\theta)}{\partial x \ \partial y} = e^{-a(\theta)b(x,y)-c(x,y)-d(\theta)}.$$

For $\theta = \theta_0 = 0$ we will suppose independence and thus $P(w|0) = 0$ and $Q(w|0) = e^{-w}$; we can suppose $a(0) = d(0) = 0$ by a simple transformation and, as $a'(0)$ is a factor of the variance, we can suppose $a'(0) = 1$.

Then, putting $\theta = 0$, we obtain $c(x,y) = x + y + e^{-x} + e^{-y}$, so that the condition for sufficiency can be written

$$(e^{-x}+e^{-y})\,k(y-x|\theta) - \log\{P\,(y-x|\theta) + e^{-x}\,Q(y-x|\theta)\}$$

$$= a(\theta)\,b(x,y) + y + e^{-x} + e^{-y} + d(\theta).$$

Denoting by $h(w) = \frac{\partial\, k(w|\theta)}{\partial\, \theta}\big|_{\theta=0}$ and deriving the relation above in order to θ, we obtain

$$(e^{-x}+e^{-y})\,h(y-x) = e^{y-x}\,\{\frac{\partial\, P(y-x|\theta)}{\partial\,\theta}\big|_{\theta=0} + e^{-x}\,\frac{\partial\, Q(y-x|\theta)}{\partial\,\theta}\big|_{\theta=0}\} + b(x,y) + d'(0)$$

so that

$$b(x,y) = e^{-x}\,\alpha(y-x) + \beta(y-x),$$

and the sufficiency condition has the form

$$(e^{-x}+e^{-y})\,k(y-x|\theta) - \log\{P\,(y-x|\theta) + e^{-x}\,Q(y-x|\theta)\} =$$

$$y + e^{-x} + e^{-y} + a(\theta)\,(e^{-x}\,\alpha(y-x) + \beta(y-x)) + d(\theta).$$

Putting $y = w + x$ and letting $x \to +\infty$ we see that we must have $P(w|\theta) = 0$ for any θ, which implies independence.

As regards correlation, we can also give some results.

The first one is that symmetrization reduces correlation.

If $\Lambda_1(x,y)$ and $\Lambda_2(x,y)$ are two distribution functions of bivariate reduced extremes with dependence functions $k_1(w)$ and $k_2(w)$ and correlation coefficients $\rho_i = -\frac{6}{\pi^2} \int_{-\infty}^{+\infty} \log k_i(w)\, d\, w$, consider now the dependence function $\bar{k}(w) = (k_i(w) + k_2(w))/2$ — also a dependence function by the mix-technique — and as $\frac{k_1+k_2}{2} \geq \sqrt{k_1\, k_2}$ we obtain

$$\bar{\rho} = -\frac{6}{\pi^2} \int_{-\infty}^{+\infty} \log \bar{k}(w)\, dw \leq -\frac{6}{\pi^2} \int_{-\infty}^{+\infty} \log \sqrt{k_1(w) k_2(w) dw} = (\rho_1 + \rho_2)/2.$$

Reference should be made to Tiago de Oliveira (1991).

Also if $k_0 \leq k(w)$, we know that $\max(k_0, \frac{\max(1,e^w)}{1+e^w}) \leq k(w)$

and thus

$$\int_{-\infty}^{+\infty} \log k(w)\, d\, w \geq \int_{-\infty}^{+\infty} \log \max(k_0, \frac{\max(1,e^w)}{1+e^w}) d\, w =$$

$$= 2 \int_0^{+\infty} \log \max(k_0, \frac{e^w}{1+e^w}) d\, w = 2\, [\int_0^{\log(k_0/(1-k_0))} \log k_0\, d\, w$$

$$+ \int_{\log(k_0/(1-k_0))}^{+\infty} \log \frac{e^w}{1+e^w}\, d\, w] = 2 \int_1^{1/k_0} \frac{\log(\xi-1)}{\xi}\, d\, \xi\,,$$

and, thus, denoting by

$$\varphi(k_0) = -\frac{12}{\pi^2} \int_1^{1/k_0} \frac{\log(\xi-1)}{\xi} d\, \xi,$$

we see that

$$\varphi(1) = 0 \leq \rho \leq \varphi(k_0) \leq \varphi(1/2) \leq= 1.$$

$\varphi(k_0)$ is a decreasing function of k_0, and the relation for k_0 can be rewritten as $1/2 \leq k_0 \leq \varphi^{-1}(\rho) \leq 1$; and as the index of dependence $\Delta(k_0)$ is also a decreasing function of k_0, we get

$$\Delta(1) = 0 \leq \Delta(\varphi^{-1}(\rho)) \leq \Delta(k_0) \leq \Delta(1/2) = 1/4$$

and so the knowledge of ρ gives a lower bound for $\Delta(k_0)$.

An inequality with the reverse direction was not obtained although such a result seems possible as we know that for $\rho = 0$ we have $k_0 = 1(k(w) = 1$, independence) and for $\rho = 1$ we have $k_0 = 1/2(k(w) = \frac{\max(1,e^w)}{1+e^w}$, diagonal case) against the fact that by the previous inequality for $\rho = 1$ we have $k_0 = 1/2$ and $\Delta(1/2) = 1$ but for $\rho = 0$ we have $k_0 = 1$ and $0 \le \Delta(k_0) \le 1$.

An important and unexpected relation between ρ and τ is

$$(1-\tau)^2 \le \frac{4\pi^2}{9}(1-\rho).$$

In fact, as we have $\rho = 1 - \frac{3}{\pi^2}\int_{-\infty}^{+\infty} w^2\, d\, D(w)$ and $\tau = 1 - \int_{-\infty}^{+\infty} w\, d\, D^2(w)$, the Schwarz inequality gives the desired result.

The relation can be written as $1-\tau \le \frac{2\pi}{3}\sqrt{1-\rho}$ or$1 - \frac{2\pi}{3}\sqrt{1-\rho} \le \tau$. Then we get

$$1 - \min(\frac{2\pi}{3}\sqrt{1-\rho}\,,1) \le \tau \le 1$$

which improves the usual $0 \le \tau \le 1$ if $\frac{2\pi}{3}\sqrt{1-\rho} \le 1$ or $\rho > 1 - (\frac{2\pi}{3})^2 = .77203$.

Also, given τ, we get $0 \le \rho \le 1 - \frac{9}{4\pi^2}(1-\tau)^2$.

Let $k_1(w)$ and $k_2(w)$ be the dependence functions and $D_1(w)$ and $D_2(w)$ the distribution functions of the reduced extremes of the distribution functions $\Lambda_1(x,y)$ and $\Lambda_2(x,y)$.

The difference between the correlation coefficients is

$$\rho_2 - \rho_1 = -\frac{6}{\pi^2}\int_{-\infty}^{+\infty} \log\frac{k_2(w)}{k_1(w)}\, d\, w.$$

Let us denote by $Q_1 = \inf\frac{k_2(w)}{k_1(w)}$ and $Q_2 = \sup\frac{k_2(w)}{k_1(w)}$; as for $w \to \pm\infty$ we know that $k_2(w)/k_1(w) \to 1$ and by the relation $k_D(w) = 1/1 + e^{-|w|} \le k_1(w), k_2(w) \le 1 = k_1(w)$, we get the inequalities $1/2 \le Q_1 \le 1 \le Q_2 \le 2$.

As we have $Q_1 k_1(w) \leq k_2(w) \leq Q_2 k_1(w)$ and by the previous relations we get $\max(k_D(w), Q_1 k_1(w)) \leq k_2(w) \leq \min(1, Q_2 k_1(w))$, then

$$-\frac{6}{\pi^2}\int_{-\infty}^{+\infty}(\log\,\min(1, Q_2 k_1(w)) - \log k_1(w))\,d\,w \leq \rho_2 - \rho_1 \leq$$

$$-\frac{6}{\pi^2}\int_{-\infty}^{+\infty}(\log\,\max(k_D(w), Q_1 k_1(w)) - \log k_1(w))\,d\,w\,;$$

$$-\frac{6}{\pi^2}\int_{-\infty}^{+\infty}(\log\,\min(Q_2, 1/k_1(w))\,d\,w \leq \rho_2 - \rho_1 \leq$$

$$-\frac{6}{\pi^2}\int_{-\infty}^{+\infty}(\log\max(\frac{k_D(w)}{k_1(w)}, Q_1))d\,w$$

with $k_D(w) \leq k_1(w), k_2(w) \leq 1$.

Then $1/k_1(w) \leq 1/k_D(w)$ and, thus, $\min(Q_2, 1/k_1(w)) \leq \min(Q_2, 1/k_D(w))$ and so

$$-\frac{6}{\pi^2}\int_{-\infty}^{+\infty}\log\min(Q_2, 1/k_D(w))\,d\,w \leq -\frac{6}{\pi^2}\int_{-\infty}^{+\infty}\log\min(\,Q_2, 1/k_1\,(w))\,d\,w.$$

But

$$\frac{6}{\pi^2}\int_{-\infty}^{+\infty}\log\min(Q_2, 1+e^{-|w|})\,d\,w = \frac{12}{\pi^2}\int_{-\infty}^{+\infty}\log\min(Q_2, 1+e^{-w})\,d\,w =$$

$$= \frac{12}{\pi^2}\,[\int_0^{-\log(Q_2-1)}\log Q_2 d\,w + \int_{-\log\,(Q_2-1)}^{+\infty}\log(1+e^{-w})\,d\,w =$$

$$= \frac{12}{\pi^2}\{-\log(Q_2-1)\,.\log Q_2 + \int_0^{Q_2-1}\frac{\log(1+t)}{t}\,d\,t]$$

$$= \frac{12}{\pi^2}\{-\log(Q_2-1)\,.\log Q_2 + \log(1+t)\log t\,|_0^{Q_2-1} - \int_0^{Q_2-1}\frac{\log t}{1+t}\,d\,t\}$$

$$= -\frac{12}{\pi^2}\int_0^{Q_2-1}\frac{\log t}{1+t}\,d\,t = -\frac{12}{\pi^2}\int_1^{Q_2}\frac{\log(\xi-1)}{\xi}\,d\,\xi = \varphi(Q_2^{-1}) \text{ and so}$$

$$-\varphi(Q_2^{-1}) \leq \rho_2 - \rho_1.$$

But also

$1 \leq 1/k_1(w))$ so $\max(\frac{k_D(w)}{k_1(w)}, Q_1) \geq \max(k_D(w), Q_1)$

and then

$$-\frac{6}{\pi^2}\int_{-\infty}^{+\infty}\log\max(\frac{k_D}{k_1},Q_1)dw \leq -\frac{6}{\pi^2}\int_{-\infty}^{+\infty}\log\max(k_D,Q_1)d\,w\,;$$

$$=\frac{6}{\pi^2}\int_{-\infty}^{+\infty}\log\min\,(\frac{1}{k_D},\frac{1}{Q_1})\;dw=\varphi(Q_1)$$

and, thus, we get

$$-\varphi(Q_2^{-1}) \leq \rho_2-\rho_1 \leq \varphi(Q_1).$$

Also, the exchange of k_1 and k_2 (exchanging Q_1 and Q_2 for Q_2^{-1} and Q_1^{-1}) gives the same result and, so, $|\rho_2-\rho_1| \leq \max(\varphi(Q_2^{-1}),\varphi(Q_1))$ and, as $\varphi(Q)$ is an increasing function for $Q^{-1}>1$, we get

$$|\rho_2-\rho_1| \leq \varphi(\min(Q^{-1},Q_2)).$$

The distance between Λ_1 and Λ_2, analogous to the index of dependence, is

$$\sup_{x,y}|\,\Lambda_1(x,y)-\Lambda_2(x,y)|=\max\{(Q_2-1)Q_2{}^{-\frac{Q_2}{Q_2-1}},(1-Q_1)Q_1{}^{\frac{Q_1}{1-Q_1}}\},$$

where Q_1 and Q_2^{-1} are analogous to k_0 in the beginning, and the technique of computation is the same. As $(Q_2-1)e^{-Q_2/(Q_2-1)} \leq Q_2-1$ and $(1-Q_1)\,Q_1{}^{Q_1/(Q_1-1)} \leq 1-Q_1$ are rough approximations, a very rough approximation is $\sup|\,\Lambda_1-\Lambda_2| \leq \max(Q_2-1,1-Q_1)$. If we consider the two non-differentiable models, biextremal with

$$k_1(w|\theta)=1-\frac{\min(\theta,e^w)}{1+e^w}$$

and its dual with

$$k_2(w|\theta)=1-\frac{\min(1,\theta\,e^w)}{1+e^w}.$$

where $0 \leq \theta \leq 1$, both with correlation coefficients

$$\rho_1(\theta)=\rho_2(\theta)=-\frac{6}{\pi^2}\int_0^\theta\frac{\log t}{1-t}\,d\,t,$$

we have

$$Q_2=\sup_w\frac{k_2(w|\theta)}{k_1(w|\theta)}=1+\theta(1-\theta),Q_1=\inf_w\frac{k_2(w|\theta)}{k_1(w|\theta)}=1/Q_2$$

so that, in this case,

$$\sup_{x,y} |\Lambda_1(x,y) - \Lambda_2(x,y)| = (Q_2 - 1)Q_2^{-\frac{Q_2}{Q_2-1}};$$

this distance is an increasing function of $Q_2 = 1 + \theta(1-\theta)$ whose maximum value is $Q_2 = 1 + 1/4$. So, with equal correlation coefficients for the biextremal model and its dual, we have the maximum distance equal to

$$\frac{1}{4} \times \left(\frac{5}{4}\right)^{-\frac{5/4}{1/4}}; = 4^4/5^5 = .08192.$$

Another expression for the difference of the correlation coefficients is

$$\rho_2 - \rho_1 = \frac{3}{\pi^2}\int_{-\infty}^{+\infty} w^2\, d(D_1(w) - D_2(w)) = \frac{6}{\pi^2}\int_{-\infty}^{+\infty} w\,(D_2(w) - D_1(w))\, d\, w.$$

Also for the difference-sign correlation we have

$$\tau_2 - \tau_1 = \frac{3}{\pi^2}\int_{-\infty}^{+\infty} w\, d(D_1^2(w) - D_2^2(w)) = \frac{3}{\pi^2}\int_{-\infty}^{+\infty} w\,(D_2^2(w) - D_1^2(w)) d\, w.$$

From the relations written between ρ and τ we get

$$|\rho_2 - \rho_1| = \frac{9}{4\pi^2}\max\{(1-\tau_1)^2, (1-\tau_2)^2 = \frac{9}{4\pi^2}(1 - \min(\tau_1,\tau_2))^2$$

and, conversely, we obtain

$$(\tau_1,\tau_2) \le \max\left(\min\left(\frac{2\pi}{3}\sqrt{1-\rho_1}, 1\right), \min\left(\frac{2\pi}{3}\sqrt{1-\rho_2}, 1\right)\right)$$

$$= \min\left(1, \frac{2\pi}{3}\max\left(\sqrt{1-\rho_1}, \sqrt{1-\rho_2}\right)\right) = \min\left(1, \frac{2\pi}{3}\sqrt{1-\min(\rho_1,\rho_2)}\right).$$

References

Berman, S. M., 1962. Limiting distribution of the maximum term in sequences of dependent random variables. *Ann. Math. Statist.*, 33, 894-908.

Cramér, H., 1946. *Mathematical Methods of Statistics*, Princeton University Press, New Jersey.

Deheuvels, P., 1984. Probabilistic aspects of multivariate extremes. *Statistical Extremes and Applications*, J. Tiago de Oliveira ed., 117-130, D. Reidel, Dordrecht.

Finkelshteyn, B. V., 1953 . Limiting distributions of extremes terms of the variational series of a two-dimensional random variable. *Dokl. Ak. Nauk. SSSR*, 91 (in Russian), 209-211.

Fraser, D. A. S., 1957. *Non -Parametric Methods in Statistics*, Wiley, New York.

Fréchet, M., 1951. Sur les tableaux de corrélation dont les marges sont données. *Ann. Univ. Lyon*, Ser III, XIV-A, Math. Et. Astr., 53-77.

Galambos, J., 1978. *The Asymptotic Theory of Extreme Order Statistics*, Wiley, New - York;(1987), 2nd. ed. Krieger Publ.

Geffroy, J., 1958/59. Contribution a l'étude de la théorie des valeurs extremes. *Publ. Inst.Statist. Univ. Paris*, 7/8, 37-185.

Gumbel, E. J. and Mustafi, C. K., 1968. Conditional medians for bivariate extreme distributions. Tech. Report. Columbia University (New York), Dep. of Industrial Engineering.

Konijn, H. S., 1956. On the power of certain tests for independence for bivariate populations. *Ann. Math. Statist.*, 27, 300-323.

Mardia, K. V., 1964. Asymptotic independence of bivariate extremes. *Calcutta Statist. Assoc. Bull.*, 13, 172-178.

Sibuya, M., 1960. Bivariate extremal statistics, I. *Ann. Inst. Statist. Math.*, XI, 195-210, Tokyo.

Tiago de Oliveira, J, 1958. Extremal distributions. *Rev. Fac. Ciências de Lisboa*, 2 ser., A, Mat, VII, 215-227.

Tiago de Oliveira, J, 1962/63. Structure theory of bivariate extremes; extensions. *Estudos Mat., Estatist. e Econometria*, Lisboa, VII, 165-194.

Tiago de Oliveira, J, 1964. L'indépendance des distributions extrêmales bivariées. *Publ.Inst. Statist. Univ. Paris*, XIII, 137-141.

Tiago de Oliveira, J, 1974. Regression in the nondifferentiable bivariate extreme models. *J Amer. Statist. Assoc*., 69, 816-818.

Tiago de Oliveira, J, 1975. Bivariate and multivariate extreme distributions. *Statistical Distributions in Scientific Work*, G. P. Patil, S. Kotz and J. K. Ord eds., 1, 355-361, D.Reidel, Dordrecht.

Tiago de Oliveira, J, 1980. Bivariate extremes: foundations and statistics. *Multivariate Analysis -V*, P. R. Krishnaiah ed., 349-368, North-Holland, Amsterdam.

Tiago de Oliveira, J, 1984. Bivariate models for extremes: statistical decision, in *Statistical Extremes and Application.* J. Tiago de Oliveira ed., 131-153, D. Reidel, Dordrecht.

Tiago de Oliveira, J, 1991. Some notes on correlation for bivariate extremes. Presented at the *48th session of the I.S.I.*, Cairo.

*****