AMSI Summer School 2023

Introduction to the modelling of continuous multivariate distributions using copulas

A very short quiz

Exercise 1

Let (X, Y) be a bivariate random vector and let H denote its distribution function (d.f.) defined by $H(x, y) = \mathbb{P}(X \leq x, Y \leq y)$, $x, y \in \mathbb{R}$.

1. Express the marginal d.f.s F and G of X and Y, respectively, in terms of H.
2. Assuming that it exists on $\mathbb{R}^2$, express the density h of (X, Y) in terms of H. In addition, express the densities f and g of X and Y, respectively, in terms of h. Finally, express H in terms of h.
3. Assume that (X, Y) follows a bivariate normal distribution. Give an expression of the densities h, f and g in that case. Can you give an easy-to-compute expression of H?

Exercise 2

Let U and V be two independent standard uniform random variables.

1. Give the expressions of the d.f. and the density of U.
2. Show that the restriction to $[0, 1]^2$ of the d.f. Π of (U, V) is given by $\Pi(u, v) = uv$, $u, v \in [0, 1]$.
3. Consider the random vectors (U, U) and $(U, 1-U)$. Use R to plot 1000 realizations of each of these random vectors.
4. Show that the restrictions to $[0, 1]^2$ of the d.f.s of these random vectors are $M(u, v) = \min\{u, v\}$ and $W(u, v) = \max\{u + v - 1, 0\}$, $u, v \in [0, 1]$, respectively. Are the two random vectors under consideration *absolutely continuous*?

The d.f.s Π, M and W are examples of *copulas*.

Solutions to exercise 1

1. With some abuse of notation, $F(x) = H(x, \infty)$, $x \in \mathbb{R}$, and $G(y) = H(\infty, y)$, $y \in \mathbb{R}$.

2. We have

$$\begin{aligned} h &= \frac{\partial^2 H}{\partial x \partial y}, \\ f(x) &= \int_{\mathbb{R}} h(x,y)\,dy, \qquad x \in \mathbb{R}, \\ g(y) &= \int_{\mathbb{R}} h(x,y)\,dx, \qquad y \in \mathbb{R}, \\ H(x,y) &= \int_{-\infty}^{x} \int_{-\infty}^{y} h(s,t)\,ds\,dt, \qquad x, y \in \mathbb{R}. \end{aligned}$$

3. See https://en.wikipedia.org/wiki/Multivariate_normal_distribution for instance. No, there is no easy-to-compute expression of H. Numerical integration must be used (as in the univariate case).

Solutions to exercise 2

1. The d.f. F of U is given by

$$F(x) = \begin{cases} 0 \text{ if } x \leq 0, \\ x \text{ if } 0 \leq x \leq 1, \\ 1 \text{ if } x \geq 1. \end{cases}$$

   The density f of U is given by

$$f(x) = \begin{cases} 0 \text{ if } x \leq 0, \\ 1 \text{ if } 0 < x < 1, \\ 0 \text{ if } x \geq 1. \end{cases}$$

2. For any $u, v \in [0, 1]$,

$$\Pi(u,v) = \mathbb{P}(U \leq u, V \leq v) = \mathbb{P}(U \leq u)\mathbb{P}(V \leq v) = uv.$$

3. Using R:

```
u <- runif(1000)
plot(u, u)
plot(u, 1 - u)
```

4. For any $u, v \in [0, 1]$,

$$M(u,v) = \mathbb{P}(U \leq u, U \leq v) = \mathbb{P}(U \leq \min\{u,v\}) = \min\{u,v\}.$$

For any $u, v \in [0,1]$,

$$W(u,v) = \mathbb{P}(U \leq u, 1-U \leq v) = \mathbb{P}(1-v \leq U \leq u) = \begin{cases} u-1+v \text{ if } u \geq 1-v, \\ 0 \text{ otherwise.} \end{cases}$$

No, these two random vectors are not absolutely continuous as their realizations are concentrated on subsets of $[0,1]^2$ (the diagonal and the anti-diagonal, respectively) of zero Lebesgue measure.