Exercise 1 (4 Points). Let (X_1, Y_1) and (X_2, Y_2) be independent vectors of absolutely continuous random variables with joint distribution functions F_{X_1,Y_1} and F_{X_2,Y_2} respectively, with common margins $F_X = F_{X_1} = F_{X_2}$ and $F_Y = F_{Y_1} = F_{Y_2}$. Let C_1 and C_2 denote the copulas of (X_1, Y_1) and (X_2, Y_2), respectively. Show that

$$P[(X_1 - X_2)(Y_1 - Y_2) > 0] - P[(X_1 - X_2)(Y_1 - Y_2) < 0] \\ = 4 \int_{[0,1]^2} C_2(u, v) dC_1(u, v) - 1 =: Q(C_1, C_2) . \quad (1)$$

In order to ensure that nothing goes wrong we follow the proof of theorem 5.1.1 from [Nel06]. Because all random variables are continuous, points have no mass so $P[(X_1 - X_2)(Y_1 - Y_2) < 0] = 1 - P[(X_1 - X_2)(Y_1 - Y_2) > 0]$ and thus it is sufficient to show that $4 \int_{[0,1]^2} C_2(u, v) dC_1(u, v) - 1 = 2P[(X_1 - X_2)(Y_1 - Y_2) > 0] - 1$ or that $P[(X_1 - X_2)(Y_1 - Y_2) > 0] = 2 \int_{[0,1]^2} C_2(u, v) dC_1(u, v)$. We have $\omega \in \{(X_1 - X_2)(Y_1 - Y_2) > 0\}$ iff either $X_1(\omega) > X_2(\omega)$ and $Y_1(\omega) > Y_2(\omega)$ or $X_1(\omega) < X_2(\omega)$ and $Y_1(\omega) < Y_2(\omega)$, so that $P[(X_1 - X_2)(Y_1 - Y_2) > 0] = P(X_1 > X_2, Y_1 > Y_2) + P(X_1 < X_2, Y_1 < Y_2)$. We will now show that each $P(X_1 > X_2, Y_1 > Y_2) = P(X_1 < X_2, Y_1 < Y_2) = \int_{[0,1]^2} C_2(u, v) dC_1(u, v)$. We have

$$P(X_1 > X_2, Y_1 > Y_2) = P(X_2 < X_1, Y_2 < Y_1) ,$$

and by definition of F_{X_1,Y_1}

$$= \int_{\mathbb{R}^2} P(X_2 \leq x, Y_2 \leq y) dF_{X_1,Y_1}(x, y) .$$

By definition of C_1 and C_2 and Sklar's theorem we get

$$= \int_{\mathbb{R}^2} C_2(F_X(x), F_Y(y)) dC_1(F_X(x), F_Y(y)) ,$$

and by choosing $u = F_X(x)$ and $v = F_Y(y)$

$$= \int_{[0,1]^2} C_2(u,v) dC_1(u,v) .$$

Analogously we can write

$$P(X_1 < X_2, Y_1 < Y_2) = \int_{\mathbb{R}^2} P(X_2 \geq x, Y_2 \geq y) dF_{X_1,Y_1}(x,y) .$$

Now since $P(A \cap B) = P(A) + P(B) - P(A \cup B) = 1 - P(A^C) - P(B^C) + P(A^C \cap B^C)$ we get

$$= \int_{\mathbb{R}^2} 1 - F_X(x) - F_Y(y) + C_2(F_X(x), F_Y(y)) .$$

again by definition of C_1 and C_2, Sklar's theorem, and by substituting $u = F_X(x)$ and $v = F_Y(y)$ we obtain

$$= \int_{[0,1]^2} 1 - u - v + C_2(u,v) dC_1(u,v) .$$

Since $C_1(u,v) = P(U \leq u, V \leq v)$ with $U \sim V \sim \mathrm{U}(0,1)$, we have $\int_{[0,1]^2} dC_1(u,v) = 1$ and $\int_{[0,1]^2} u dC_1(u,v) = \int_{[0,1]^2} v dC_1(u,v) = \frac{1}{2}$, so that also in this case we end up with

$$= \int_{[0,1]^2} C_2(u,v) dC_1(u,v) .$$

Exercise 2 (4 Points). Let (X, Y) be a vector of absolutely continuous random variables. Use the definition of Q in (1) to show that

1. the population version of Kendall's τ is given by $\tau(X, Y) = Q(C, C)$ if C is a copula for (X, Y) and

   this can be directly seen from (1) by setting $C_1 = C_2 = C$.

2. the population version of Spearman's ρ is given by $\rho(X, Y) = 3Q(C, \Pi)$, if C is a copula for (X, Y) and Π describes the independence copula $\Pi(u, v) = uv$.

   Spearman's ρ is given by

$$\frac{1}{3}\rho(X, Y) = P[(X_1 - X_2)(Y_1 - Y_3) > 0] - P[(X_1 - X_2)(Y_1 - Y_3) < 0]\,,$$

where (X_1, Y_1), (X_2, Y_2) and (X_3, Y_3) are three independent vectors with joint distribution function $F_{X,Y}$. In analogy to exercise 1 we can write

$$= 4 \int_{[0,1]^2} C_2(u, v) dC_1(u, v) - 1\,,$$

where $C_1 = C$ is again the copula of (X_1, Y_1), but C_2 is now the copula of (X_2, Y_3), because in the formula now there is Y_3 where Y_2 was. Since (X_2, Y_2) and (X_3, Y_3) are independent, $C_2 = \Pi$ so that we get

$$= Q(C, \Pi)\,.$$

Exercise 3 (4 Points). Show that $3Q(C,\Pi) = 12\int_0^1\int_0^1 (C(u,v) - uv)dudv$.

Since Q is symmetric in its arguments, *which still remains to be shown,* we can equivalently write $\rho(X,Y) = 12\int_0^1\int_0^1 C(u,v)dudv - 3$. The result then follows $\int_0^1\int_0^1 uvdudv = \frac{1}{4}$.

Then show that Spearman's ρ is a measure of concordance.

To this end we have to show that

1. ρ is defined for every pair X, Y of continuous random variables

which is true by definition.

2.1. $-1 \leq \rho \leq 1$

which is true, because $0 \leq P[(X_1 - X_2)(Y_1 - Y_3) > 0] \leq 1$ and $0 \leq P[(X_1 - X_2)(Y_1 - Y_3) < 0] \leq 1$.

2.2. $\rho(X,X) = 1$

which is true, because then $C(u,v) = \min(u,v)$ so that $3Q(C,\Pi) = 12\int_0^1\int_0^1(\min(u,v) - uv)dudv = 1$.

3. $\rho(X,Y) = \rho(Y,X)$

which follows from Q being symmetric in its arguments.

4. If X and Y are independent, then $\rho(X,Y) = 0$

which follows from the fact that then $C(u,v) = uv$ so that $C(u,v) - uv = 0$.

6. If C_1 and C_2 are copulas for (X_1,Y_1) and (X_2,Y_2) such that $C_1 \prec C_2$ then $\rho(X_1,Y_1) \leq \rho(X_2,Y_2)$,

which follows from the monotony of the integral.

7. If $\{(X_n,Y_n)\}$ is a sequence of continuous random variables with copulas C_n, and if $\{C_n\}$ converges pointwise to C, the copula for (X,Y) then $\lim_{n\to\infty}\rho(X_n,Y_n) = \rho(X,Y)$.

Because all the copulas are quasi-copulas, they fulfill $|C_n(u,v)| \leq |u| + |v|$ and $|C(u,v)| \leq |u| + |v|$, so that for all $\varepsilon > 0$ we can find an $n \in \mathbb{N}$ so that $|C_n(u,v) - C(u,v)| < \varepsilon$, *where however I don't know how.* That means C_n is uniformly convergent towards C and we get $\lim_{n\to\infty} \int_0^1 \int_0^1 C_n(u,v)dudv = \int_0^1 \int_0^1 C(u,v)dudv$ as desired.

References

[Nel06] Nelsen, Roger B.: *An introduction to copulas.* Springer, 2006