Lecture Notes on Copula

Giovanni Della Lunga

April 2018

1 Introduction

Many real-life situations can be modelled by a large number of random variables which play a significant role, and such variates are generally not independent. Therefore, it is often of fundamental importance to be able to link the marginal distributions of different variables in order to give a flexible and accurate descrip- tion of the joint law of the variables of interest. Copulas were introduced in 1959 in the context of probabilistic metric spaces and later exploited as a tool for understanding relationships among multivariate outcomes.

A copula is a function that links univariate marginals to their joint multivariate distribution in such a way that it captures the entire dependence structure in the multivariate distribution. The main advantage provided by a copula-approach in dependence modelling is that the selection of an appropriate model for the dependence between variables X and Y , represented by the copula, can proceed independently from the choice of the marginal distributions. The seminal result in the history of copulas is due to Sklar that introduced in 1959 the notion, and the name, of copula, and proved the theo- rem that now bears his name (Sklar, 1959). The latter states that any multivariate distribution can be expressed as its copula function evaluated at its marginal distribution functions. Moreover, any copula function when evaluated at any marginal distributions is a multivariate distribution.

2 Joint Cumulative Distribution Function

Remember that the ¡b¿joint cumulative function¡/b¿ of two random variables X and Y is defined as

$$F_{XY}(x,y) = P(X \leq x, Y \leq y).$$

The joint CDF satisfies the following properties:

1. $F_X(x) = F_{XY}(x,\infty)$, for any x (marginal CDF of X);
2. $F_Y(y) = F_{XY}(\infty,y)$, for any y (marginal CDF of Y);
3. $F_{XY}(\infty,\infty) = 1$;
4. $F_{XY}(-\infty,y) = F_{XY}(x,-\infty) = 0$;
5. $\mathbb{P}(x_1 < X \leq x_2,\ \ y_1 < Y \leq y_2) = F_{XY}(x_2,y_2) - F_{XY}(x_1,y_2) - F_{XY}(x_2,y_1) + F_{XY}(x_1,y_1)$;
6. if X and Y are independent, then $F_{XY}(x,y) = F_X(x)F_Y(y)$.

In particular from property 5, putting $x_2 \to +\infty$ and $y_2 \to +\infty$ we have

$$\begin{aligned} \mathbb{P}(x_1 < X \leq +\infty,\ \ y_1 < Y \leq +\infty) &= F_{XY}(+\infty,+\infty) - F_{XY}(x_1,+\infty) - F_{XY}(+\infty,y_1) + F_{XY}(x_1,y_1) && (1) \\ &= 1 - F_X(x) - F_Y(y) + F_{XY}(x_1,y_1) && (2) \end{aligned}$$

If we denote with

$$\bar{F}_{XY}(x,y) = \mathbb{P}[X > x, Y > y] \tag{3}$$

we finally obtain

$$\bar{F}_{XY}(x,y) = 1 - F_X(x) - F_Y(y) + F_{XY}(x,y) \tag{4}$$

3 Survival Copula

The survival copula associated with the copula C is

$$\bar{C}(u,v) = u + v - 1 + C(1-u, 1-v) \tag{5}$$

It is easy to verify that $\bar{C}$ ha the copula properties. Once computed in $(1-u, 1-v)$ is equivalent to the complementary distribution function of a bivariate uniform distribution, since

$$\begin{aligned}
\bar{C}(1-u,1-v) &= 1 - u + 1 - v - 1 + C(u,v) && (6)\\
&= 1 - \mathbb{P}(U_1 \leq u) + 1 - \mathbb{P}(U_2 \leq v) - 1 + \mathbb{P}(U_1 \leq u, U_2 \leq v) && (7)\\
&= 1 - \mathbb{P}(U_1 \leq u) - \mathbb{P}(U_2 \leq v) + \mathbb{P}(U_1 \leq u, U_2 \leq v) && (8)\\
&= \mathbb{P}(U_1 > u, U_2 > v) && (9)
\end{aligned}$$

4 Copula and Measure

If C is absolutely continuous then it can be written in the form

$$C(\mathbf{u}) = \int\limits_{[0,\mathbf{u}]^d} c(\mathbf{s})\, d\mathbf{s} \tag{10}$$

where c is a suitable function called density of C. In particular, for almost all $\mathbf{u} \in \mathbb{I}^d$ one has

$$c(\mathbf{u}) = \frac{\partial^d C(\mathbf{u})}{\partial u_1 \ldots \partial u_d} \tag{11}$$

As stressed by many authors, this equation is far from obvious. In fact, there are some facts that are implicitly used: first, the mixed partial derivatives of order d of C exist and are equal almost everywhere on $\mathbb{I}^d$; second each mixed partial derivative is actually almost everywhere equal to the density c. We will use this result to formally define a measure induced on $\mathbb{I}^d$ by C.

$$dC(\mathbf{u}) = c(\mathbf{u})\, d\mathbf{u} \tag{12}$$

In particular for a bivariate copula we have

$$dC(u,v) = \frac{\partial^2 C(u,v)}{\partial u \partial v} du dv \tag{13}$$

5 Density Function for the Sum of Correlated Gaussian Variables

When two random variables are independent, the probability density function for their sum is the convolution of the density functions for the variables that are summed. We consider here the case when these two random variables are correlated. Let x and y be the two correlated random variables, and $z = x + y$. The standard procedure for obtaining the distribution of a function $z = g(x,y)$ is to integrate the joint density function $p_{xy}(x,y)$ over the region D of the xy plane where $g(x,y) < Z$ to obtain the cumulative distribution $P_z(Z)$. This is then differentiated with respect to Z to obtain the density function $p_z(z)$. The cumulative distribution is:

$$P_z(Z) = \iint\limits_D p_{xy}(x,y)\, dx\, dy \tag{14}$$

In some cases, D may be disjoint, but for $z = x + y$ it is just the area under that line. When x and y are independent, the joint density function separates into a product of the two marginal density functions $p_x(x)$ and $p_y(y)$, and the procedure we are about to describe, along with $y = z - x$, leads directly to the convolution:

$$\begin{aligned}
P_z(Z) &= \int\limits_{-\infty}^{+\infty} \left(\int\limits_{-\infty}^{Z-x} p_x(x) p_y(y)\, dy \right) dx \\
p_Z(z) &= \frac{dP_z(Z)}{dZ} = \int\limits_{-\infty}^{+\infty} p_x(x) p_y(z-x)\, dx
\end{aligned} \tag{15}$$

Since we are considering x and y to be correlated, we must retain the joint density function itself:

$$
\begin{aligned}
P_z(Z) &= \int_{-\infty}^{+\infty} \left(\int_{-\infty}^{Z-x} p_{xy}(x,y)\, dy \right) dx \\
p_Z(z) &= \frac{dP_z(Z)}{dZ} = \int_{-\infty}^{+\infty} p_{xy}(x, z-x)\, dx
\end{aligned}
\tag{16}
$$

using the Leibnitz rule for the derivation of integrals.

The failure of the joint density function to separate results in $p_z(z)$ no longer having the form of a convolution. To illustrate this we consider the Gaussian case. The joint density function for Gaussian x and y coupled through a correlation coefficient ρ is

$$
p(x,y) = \frac{\exp\left[-\frac{1}{2(1-\rho^2)}\left(\frac{(x-\bar{x})^2}{\sigma_x^2} + \frac{(y-\bar{y})^2}{\sigma_y^2} - \frac{2\rho(x-\bar{x})(y-\bar{y})}{\sigma_x\sigma_y}\right)\right]}{2\pi\sigma_x\sigma_y\sqrt{1-\rho^2}}
\tag{17}
$$

where we dont't assume that either x or y is zero-mean. The computation of the integral is extremely tedious but trivial. First of all let's re-write the argument of the exponential function using the following change of variables:

$$
z = x + y \Rightarrow y = z - x, \quad \bar{y} = \bar{z} - \bar{x} \Rightarrow (y - \bar{y}) = (z - \bar{z}) - (x - \bar{x})
\tag{18}
$$

and we obtain

$$
\begin{aligned}
&\frac{(x-\bar{x})^2}{\sigma_x^2} + \frac{(y-\bar{y})^2}{\sigma_y^2} - \frac{2\rho(x-\bar{x})(y-\bar{y})}{\sigma_x\sigma_y} \\
&= \frac{\sigma_y^2(x-\bar{x})^2 + \sigma_x^2[(z-\bar{z})-(x-\bar{x})]^2 - 2\rho\sigma_x\sigma_y(x-\bar{x})[(z-\bar{z})-(x-\bar{x})]}{\sigma_x^2\sigma_y^2} \\
&= \frac{1}{\sigma_x^2\sigma_y^2}\Big\{\sigma_y^2(x-\bar{x})^2 + \sigma_x^2(z-\bar{z})^2 + \sigma_x^2(x-\bar{x})^2 - 2\sigma_x^2(z-\bar{z})(x-\bar{x}) - \\
&- 2\rho\sigma_x\sigma_y(z-\bar{z})(x-\bar{x}) + 2\rho\sigma_x\sigma_y(x-\bar{x})^2\Big\} \\
&= \frac{1}{\sigma_x^2\sigma_y^2}\Big\{\sigma_z^2(x-\bar{x})^2 + \sigma_x^2(z-\bar{z})^2 - 2(z-\bar{z})(x-\bar{x})(\sigma_x^2 + \rho\sigma_x\sigma_y)\Big\}
\end{aligned}
\tag{19}
$$

where

$$
\sigma_z^2 = \sigma_x^2 + \sigma_y^2 + 2\rho\sigma_x\sigma_y
\tag{20}
$$

$$
\tilde{x} = x - \bar{x}, \quad \tilde{z} = z - \bar{z}
\tag{21}
$$

performing the algebraic computation and rearranging terms inside braces we obtain

$$
\frac{1}{\sigma_x^2\sigma_y^2}\Big\{\sigma_z^2\tilde{x}^2 - 2\tilde{x}\tilde{z}(\sigma_x^2 + \rho\sigma_x\sigma_y) + \sigma_x^2\tilde{z}^2\Big\}
\tag{22}
$$

Put

$$
\alpha = \left[\frac{\sigma_x^2 + \rho\sigma_x\sigma_y}{\sigma_z}\right]
\tag{23}
$$

and re-write the previous term adding and subtracting the factor $\alpha^2\tilde{z}^2$ (for simplicity we don't take into account for the moment the multiplicative factor $1/\sigma_x^2\sigma_y^2$)

$$
\begin{aligned}
\sigma_z^2\tilde{x}^2 - 2\tilde{x}\tilde{z}(\sigma_x^2 + \rho\sigma_x\sigma_y) + \sigma_x^2\tilde{z}^2 &= \sigma_z^2\tilde{x}^2 - 2\tilde{x}\tilde{z}(\sigma_x^2 + \rho\sigma_x\sigma_y) + \sigma_x^2\tilde{z}^2 + \alpha^2\tilde{z}^2 - \alpha^2\tilde{z}^2 \\
&= \left[\sigma_z^2\tilde{x}^2 - 2\tilde{x}\tilde{z}(\sigma_x^2 + \rho\sigma_x\sigma_y) + \frac{\left(\sigma_x^2 + \rho\sigma_x\sigma_y\right)^2}{\sigma_z^2}\tilde{z}^2\right] + \sigma_x^2\tilde{z}^2 - \alpha^2\tilde{z}^2 \\
&= \Big[\sigma_z\tilde{x} - \alpha\tilde{z}\Big]^2 + \Big[\sigma_x^2 - \alpha^2\Big]\tilde{z}^2 \\
&= \sigma_z^2\Big[\tilde{x} - \frac{\alpha}{\sigma_z}\tilde{z}\Big]^2 + \Big[\sigma_x^2 - \alpha^2\Big]\tilde{z}^2
\end{aligned}
\tag{24}
$$

Let's compute the second term

$$
\begin{aligned}
(\sigma_x^2 - \alpha^2)\tilde{z}^2 &= \left[\sigma_x^2 - \frac{\sigma_x^4 + \rho^2\sigma_x^2\sigma_y^2 + 2\rho\sigma_x^3\sigma_y}{\sigma_z^2}\right]\tilde{z}^2 \\
&= \left[\frac{\sigma_x^2\sigma_z^2 - \sigma_x^4 - \rho^2\sigma_x^2\sigma_y^2 - 2\rho\sigma_x^3\sigma_y}{\sigma_z^2}\right]\tilde{z}^2 \\
&= \frac{1}{\sigma_z^2}\left[(\sigma_x^2 + \sigma_y^2 + 2\rho\sigma_x\sigma_y)\sigma_x^2 - \sigma_x^4 - \rho^2\sigma_x^2\sigma_y^2 - 2\rho\sigma_x^3\sigma_y\right]\tilde{z}^2 \\
&= \frac{1}{\sigma_z^2}\left[\sigma_x^4 + \sigma_y^2\sigma_x^2 + 2\rho\sigma_x^3\sigma_y - \sigma_x^4 - \rho^2\sigma_x^2\sigma_y^2 - 2\rho\sigma_x^3\sigma_y\right]\tilde{z}^2 \\
&= \frac{1}{\sigma_z^2}\left[(1-\rho^2)\sigma_x^2\sigma_y^2\right]\tilde{z}^2
\end{aligned} \tag{25}
$$

This give us

$$
\sigma_z^2\tilde{x}^2 - 2\tilde{x}\tilde{z}(\sigma_x^2 + \rho\sigma_x\sigma_y) + \sigma_x^2\tilde{z}^2 = \sigma_z^2\left[\tilde{x} - \frac{\alpha}{\sigma_z}\tilde{z}\right]^2 + \frac{1}{\sigma_z^2}\left[(1-\rho^2)\sigma_x^2\sigma_y^2\right]\tilde{z}^2 \tag{26}
$$

Remember that we have to multiply by

$$
-\frac{1}{2(1-\rho^2)\sigma_x^2\sigma_y^2} \tag{27}
$$

So, finally, the argument of the exponential function can be written as

$$
-\frac{\sigma_z^2}{2(1-\rho^2)\sigma_x^2\sigma_y^2}\left(\tilde{x} - \frac{\alpha}{\sigma_z}\tilde{z}\right)^2 - \frac{1}{2\sigma_z^2}\tilde{z}^2 \tag{28}
$$

and the integral can be written as

$$
\begin{aligned}
p_Z(z) &= \int_{-\infty}^{+\infty} p_{xy}(x, z-x)\,dx \\
&= \frac{1}{2\pi\sigma_x\sigma_y\sqrt{1-\rho^2}} \int_{-\infty}^{+\infty} \exp\left[-\frac{\sigma_z^2}{2(1-\rho^2)\sigma_x^2\sigma_y^2}\left(\tilde{x} - \frac{\alpha}{\sigma_z}\tilde{z}\right)^2 - \frac{1}{2\sigma_z^2}\tilde{z}^2\right] dx \\
&= \frac{1}{2\pi\sigma_x\sigma_y\sqrt{1-\rho^2}} \exp\left[-\frac{1}{2\sigma_z^2}\tilde{z}^2\right] \int_{-\infty}^{+\infty} \exp\left[-\frac{\sigma_z^2}{2(1-\rho^2)\sigma_x^2\sigma_y^2}\left(\tilde{x} - \frac{\alpha}{\sigma_z}\tilde{z}\right)^2\right] dx
\end{aligned} \tag{29}
$$

using

$$
\begin{aligned}
\mu &= \frac{\alpha}{\sigma_z}\tilde{z} \\
\tilde{\sigma}^2 &= (1-\rho^2)\frac{\sigma_x^2\sigma_y^2}{\sigma_z^2}
\end{aligned} \tag{30}
$$

and

$$
\tilde{\omega} = \frac{\tilde{x} - \mu}{\tilde{\sigma}} \Rightarrow d\tilde{\omega} = \frac{1}{\tilde{\sigma}}d\tilde{x} = \frac{1}{\tilde{\sigma}}dx \tag{31}
$$

we can re-write the integral as

$$
\int_{-\infty}^{+\infty} \exp\left[-\frac{\sigma_z^2}{2(1-\rho^2)\sigma_x^2\sigma_y^2}\left(\tilde{x} - \frac{\alpha}{\sigma_z}\tilde{z}\right)^2\right] dx = \tilde{\sigma}\int_{-\infty}^{+\infty} \exp\left(-\frac{\omega^2}{2}\right) dw = \sqrt{2\pi}\tilde{\sigma} \tag{32}
$$

and

$$
\begin{aligned}
p_Z(z) &= \int\limits_{-\infty}^{+\infty} p_{xy}(x, z-x)\, dx \\
&= \frac{1}{2\pi\sigma_x\sigma_y\sqrt{1-\rho^2}} \exp\left[-\frac{1}{2\sigma_z^2}\tilde{z}^2\right] \int\limits_{-\infty}^{+\infty} \exp\left[-\frac{\sigma_z^2}{2(1-\rho^2)\sigma_x^2\sigma_y^2}\left(\tilde{x}-\frac{\alpha}{\sigma_z}\tilde{z}\right)^2\right] dx \\
&= \frac{1}{2\pi\sigma_x\sigma_y\sqrt{1-\rho^2}} \exp\left[-\frac{1}{2\sigma_z^2}\tilde{z}^2\right] \left(\sqrt{1-\rho^2}\right) \frac{\sigma_x\sigma_y}{\sigma_z}\sqrt{2\pi} \\
&= \frac{1}{\sqrt{2\pi}\sqrt{\sigma_x^2+\sigma_y^2+2\rho\sigma_x\sigma_y}} \exp\left[-\frac{1}{2\sigma_z^2}\tilde{z}^2\right] \\
&= \frac{1}{\sqrt{2\pi(\sigma_x^2+\sigma_y^2+2\rho\sigma_x\sigma_y)}} \exp\left[-\frac{(z-(\bar{x}+\bar{y}))^2}{2(\sigma_x^2+\sigma_y^2+2\rho\sigma_x\sigma_y)}\right]
\end{aligned} \tag{33}
$$

which is gaussian with

$$
\begin{aligned}
\bar{z} &= \bar{x} + \bar{y} \\
\sigma_z^2 &= \sigma_x^2 + \sigma_y^2 + 2\rho\sigma_x\sigma_y
\end{aligned} \tag{34}
$$

So the mean is unaffected by the correlation, but the variance is made larger or smaller according to whether the correlation is positive or negative, respectively.

6 On Hoeffding's Identity

The Hoeffding's Identity can be written as

$$
\operatorname{cov}(X,Y) = \int_{\mathbb{R}\times\mathbb{R}} [F_{XY}(x,y) - F_X(x)F_Y(y)]\, dx\, dy \tag{35}
$$

It is interesting, and quite important, actually, to understand that the covariance (or the correlation) can be seen as some 'distance' to the independence. More precisely, observe that

$$
\operatorname{cov}(X,Y) = \int_{\mathbb{R}\times\mathbb{R}} [F_{XY}(x,y) - F^{\perp}_{X,Y}(x,y)]\, dx\, dy \tag{36}
$$

where $F^{\perp}_{X,Y}(x,y)$ would be the joint cumulative distribution function of some independent variables, with the same marginal distributions.

Now, the thing is that the proof is not trivial! Let I denote the indicator function, then we can write

$$
\begin{aligned}
\int\limits_0^{\infty} [I(z\le u) - I(x\le u)]du &= \lim_{\alpha\to\infty} \int\limits_0^{\alpha} [I(z\le u) - I(x\le u)]du \\
&= \lim_{\alpha\to\infty} \Big\{ \int\limits_0^{z} I(z\le u)\, du + \int\limits_z^{\alpha} I(z\le u)\, du \\
&\quad - \int\limits_0^{x} I(x\le u)\, du - \int\limits_x^{\alpha} I(x\le u)\, du \Big\} \\
&= \lim_{\alpha\to\infty} \Big\{ \int\limits_z^{\alpha} 1\, du - \int\limits_x^{\alpha} 1\, du \Big\} \\
&= \lim_{\alpha\to\infty} [\alpha - z - \alpha + x] = x - z
\end{aligned} \tag{37}
$$

Hence for $X_1, Y_1, X_2, Y_2 \geq 0$ we find

$$
\begin{aligned}
(X_1 - X_2)(Y_1 - Y_2) = \int_{\mathbb{R}\times\mathbb{R}} \Big[& I(X_2 \le u)I(Y_2 \le v) + I(X_1 \le u)I(Y_1 \le v) \\
& - I(X_2 \le u)I(Y_1 \le v) - I(X_1 \le u)I(Y_2 \le v)\Big]\, du\, dv
\end{aligned}
\tag{38}
$$

Now take the expectation of both sides and invert the integral and the expectation in the right hand side of the equation

$$
\begin{aligned}
\mathbb{E}[(X_1 - X_2)(Y_1 - Y_2)] = \int_{\mathbb{R}\times\mathbb{R}} \Big[& \mathbb{E}(I(X_2 \le u)I(Y_2 \le v)) + \mathbb{E}(I(X_1 \le u)I(Y_1 \le v)) \\
& - \mathbb{E}(I(X_2 \le u)I(Y_1 \le v)) - \mathbb{E}(I(X_1 \le u)I(Y_2 \le v))\Big]\, du\, dv
\end{aligned}
\tag{39}
$$

Let (X_1, Y_1) and (X_2, Y_2) be independent identically distributed pairs, using the fact that

$$
\mathbb{E}[I(X \le u)] = \mathbb{P}(X \le u) \tag{40}
$$

we obtain

$$
\begin{aligned}
\mathbb{E}[(X_1 - X_2)(Y_1 - Y_2)] &= \int_{\mathbb{R}\times\mathbb{R}} \Big[\mathbb{P}(X_2 \le u, Y_2 \le v) + \mathbb{P}(X_1 \le u, Y_1 \le v) \\
&\quad - \mathbb{P}(X_2 \le u)\mathbb{P}(Y_1 \le v) - \mathbb{P}(X_1 \le u)\mathbb{P}(Y_2 \le v)\Big]\, du\, dv \\
&= 2\int_{\mathbb{R}\times\mathbb{R}} \Big[\mathbb{P}(X_1 \le u, Y_1 \le v) - \mathbb{P}(X_1 \le u)\mathbb{P}(Y_1 \le v)\Big]\, du\, dv \\
&= 2\int_{\mathbb{R}\times\mathbb{R}} \Big[F_{X_1,Y_1}(u,v) - F_{X_1}(u)F_{Y_1}(v)\Big]\, du\, dv
\end{aligned}
\tag{41}
$$

but

$$
\begin{aligned}
\mathbb{E}[(X_1 - X_2)(Y_1 - Y_2)] &= \mathbb{E}[X_1Y_1 - X_1Y_2 - X_2Y_1 + X_2Y_2] \\
&= \mathbb{E}[X_1Y_1] - \mathbb{E}[X_1]\mathbb{E}[Y_2] - \mathbb{E}[X_2]\mathbb{E}[Y_1] + \mathbb{E}[X_2Y_2] \\
&= 2\Big(\mathbb{E}[X_1Y_1] - \mathbb{E}[X_1]\mathbb{E}[Y_1]\Big) \\
&= 2\,\mathrm{cov}(X_1, Y_1)
\end{aligned}
\tag{42}
$$

So we get Hoeffding's identity. This identity can be written using the definition of copula as

$$
\mathrm{cov}(X,Y) = \int_{\mathbb{R}\times\mathbb{R}} [C(F_X(x), F_Y(y)) - F_X(x)F_Y(y)]\, dx\, dy \tag{43}
$$

and taking into account the Fréchet inequality:

$$
\begin{aligned}
\int_{\mathbb{R}\times\mathbb{R}} [C^-(F_X(x), F_Y(y)) - F_X(x)F_Y(y)]\, dx\, dy \le \mathrm{cov}(X,Y) \\
\mathrm{cov}(X,Y) \le \int_{\mathbb{R}\times\mathbb{R}} [C^+(F_X(x), F_Y(y)) - F_X(x)F_Y(y)]\, dx\, dy
\end{aligned}
\tag{44}
$$

where C^- is the **minimum copula** and C^+ is the **maximum copula**.

7 Dependence Concepts

Copulas provide a natural way to study and measure dependence between random variables. Copula properties are invariant under strictly increasing transformations of the underlying random variables. Linear correlation (or Pearson's correlation) is most frequently used in practice as a measure of dependence. However, since linear correlation is not a copula-based measure of dependence, it can often be quite misleading and should not be taken as the canonical dependence measure. Below we recall the basic properties of linear correlation, and then continue with some copula based measures of dependence.

7.1 Linear Correlation Coefficient

Definition. Let $(X, Y)^T$ be a vector of random variables with nonzero finite variances. The linear correlation coefficient for $(X, Y)^T$ is

$$\rho(X, Y) = \frac{Cov(X, Y}{\sqrt{Var(X)}\sqrt{Var(Y)}} \tag{45}$$

where $Cov(X, Y) = \mathbb{E}(XY) - \mathbb{E}(X)\mathbb{E}(Y)$ is the covariance of $(X, Y)^T$, and $Var(X)$ and $Var(Y)$ are the variances of X and Y.

Linear correlation is a popular but also often misunderstood measure of dependence. The popularity of linear correlation stems from the ease with which it can be calculated and it is a natural scalar measure of dependence in elliptical distributions (with well known members such as the multivariate normal and the multivariate t-distribution). However most random variables are not jointly elliptically distributed, and using linear correlation as a measure of dependence in such situations might prove very misleading. Even for jointly elliptically distributed random variables there are situations where using linear correlation, as defined by (??), does not make sense. We might choose to model some scenario using heavy-tailed distributions such as t2-distributions. In such cases the linear correlation coefficient is not even defined because of infinite second moments.

Property. $\rho(X, Y)$ is bounded:

$$\rho_l \leq \rho \leq \rho_u$$

where the bounds ρ_l and ρ_u are defined as

$$\begin{aligned}
\rho_l &= \frac{\iint_D \left[C^-(F_1(x), F_2(y)) - F_1(x)F_2(y)\right] dx\, dy}{\sqrt{\left[\int_{Dom(F_1)}(x - \mathbb{E}(x))^2 dF_1(x)\right]\left[\int_{Dom(F_2)}(y - \mathbb{E}(y))^2 dF_2(y)\right]}} \\
\rho_u &= \frac{\iint_D \left[C^+(F_1(x), F_2(y)) - F_1(x)F_2(y)\right] dx\, dy}{\sqrt{\left[\int_{Dom(F_1)}(x - \mathbb{E}(x))^2 dF_1(x)\right]\left[\int_{Dom(F_2)}(y - \mathbb{E}(y))^2 dF_2(y)\right]}}
\end{aligned} \tag{46}$$

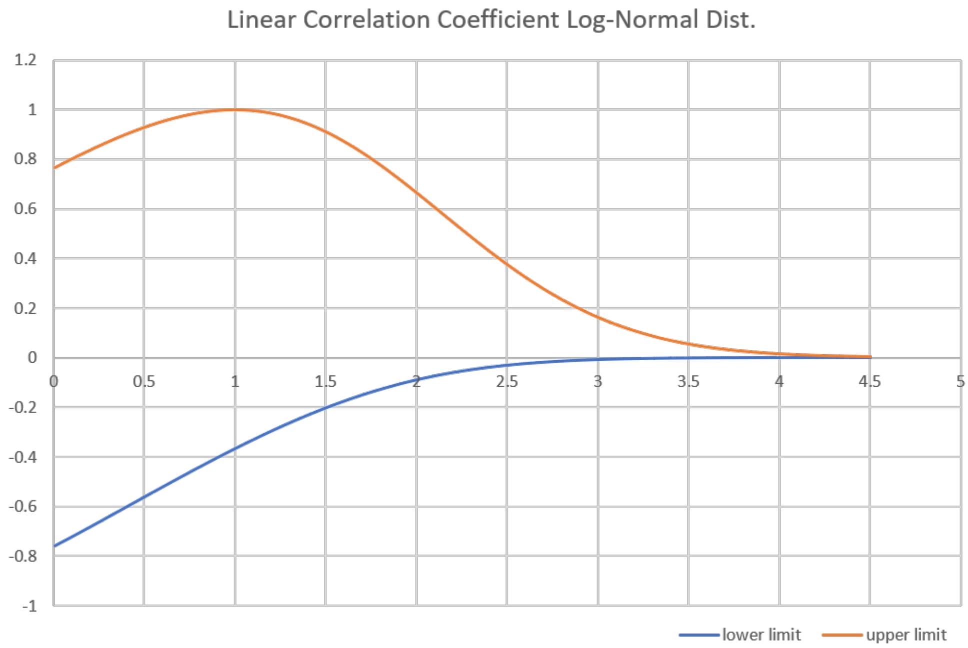
and are attained respectively when X and Y are countermonotonic and comonotonic.
Proof. The bound for $\rho(X, Y)$ can be obtained directly from Hoeffding's expression for covariance together with the Fréchet inequality.

Example. Let $X \sim LN(0, \sigma_1^2)$ and $Y \sim LN(0, \sigma_2^2)$. Then $\rho_{min} = \rho(e^{\sigma_1 Z}, e^{-\sigma_2 Z})$ and $\rho_{max} = \rho(e^{\sigma_1 Z}, e^{\sigma_2 Z})$, where $Z \sim N(0, 1)$. ρ_{min} and ρ_{max} can be computed (see Exercise ...) yielding:

$$\begin{aligned}
\rho_l &= \frac{exp(-\sigma_1\sigma_2) - 1}{\sqrt{(exp(\sigma_1^2) - 1)}\sqrt{(exp(\sigma_2^2) - 1)}} \leq 0 \\
\rho_u &= \frac{exp(\sigma_1\sigma_2) - 1}{\sqrt{(exp(\sigma_1^2) - 1)}\sqrt{(exp(\sigma_2^2) - 1)}} \geq 0
\end{aligned} \tag{47}$$

from which follows that

$$\lim_{\sigma\to\infty} \rho_{min} = \lim_{\sigma\to\infty} \rho_{max} = 0$$



7.2 Concordance

Concordance concepts, loosely speaking, aim at capturing the fact that the probability of having "large" (or "small") values of both X and Y is high, while the probability of having "large" values of X together with "small" values of "Y" - or viceversa - is low.

Let $(x, y)^T$ and $(\tilde{x}, \tilde{y})^T$ be two observations from a vector $(X, Y)^T$ of continuous random variables. Then $(x, y)^T$ and $(\tilde{x}, \tilde{y})^T$ are said to be concordant if $(x - \tilde{x})(y - \tilde{y}) > 0$, and discordant if $(x - \tilde{x})(y - \tilde{y}) < 0$.

The following theorem can be found in Nelsen (1999) p. 127. Many of the results in this section are direct consequences of this theorem.

Theorem. Let $(X, Y)^T$ and $(\tilde{X}, \tilde{Y})^T$ be independent vectors of continuous random variables with joint distribution functions H and $\tilde{H}$, respectively, with common margins F (of X and $\tilde{X}$) and G (of Y and $\tilde{Y}$). Let C and $\tilde{C}$ denote the copulas of $(X, Y)^T$ and $(\tilde{X}, \tilde{Y})^T$ respectively,so that $H(x, y) = C(F(x), G(y))$ and $\tilde{H}(x, y) = \tilde{C}(F(x), G(y))$. Let Q denote the difference between the probability of concordance and discordance of $(X, Y)^T$ and $(\tilde{X}, \tilde{Y})^T$, i.e. let

$$Q = \mathbb{P}[(X - \tilde{X})(Y - \tilde{Y}) > 0] - \mathbb{P}[(X - \tilde{X})(Y - \tilde{Y}) < 0] \tag{48}$$

Then

$$Q = Q(C, \tilde{C}) = 4 \iint_{[0,1]^2} \tilde{C}(u, v) dC(u, v) - 1 \tag{49}$$

Proof. Since the random variables are all continuous,

$$\mathbb{P}[(X - \tilde{X})(Y - \tilde{Y}) < 0] = 1 - \mathbb{P}[(X - \tilde{X})(Y - \tilde{Y}) > 0] \Rightarrow Q = 2\mathbb{P}[(X - \tilde{X})(Y - \tilde{Y}) > 0] - 1$$

But

$$\mathbb{P}[(X - \tilde{X})(Y - \tilde{Y}) > 0] = \mathbb{P}[X > \tilde{X}, Y > \tilde{Y}] + \mathbb{P}[X < \tilde{X}, Y < \tilde{Y}]$$

and these probabilities can be evaluated by integrating over the distribution of one of the vectors $(X, Y)^T$ or $(\tilde{X}, \tilde{Y})^T$. Hence

$$\begin{aligned}
\mathbb{P}[X > \tilde{X}, Y > \tilde{Y}] &= \mathbb{P}[\tilde{X} < X, \tilde{Y} < Y] && (50)\\
&= \iint_{\mathbb{R}^2} \mathbb{P}[\tilde{X} < x, \tilde{Y} < y] \, dC[F(x), G(y)] && (51)\\
&= \iint_{\mathbb{R}^2} \tilde{C}[F(x), G(y)] \, dC[F(x), G(y)] && (52)
\end{aligned}$$

Employing the probability-integral transform $u = F(x)$ and $v = G(y)$ then yields

$$\mathbb{P}[X > \tilde{X}, Y > \tilde{Y}] = \mathbb{P}[\tilde{X} < X, \tilde{Y} < Y] = \iint\limits_{[0,1]^2} \tilde{C}(u,v)\, dC(u,v) \tag{53}$$

Similarly,

$$\mathbb{P}[X < \tilde{X}, Y < \tilde{Y}] = \iint\limits_{\mathbb{R}^2} \mathbb{P}[\tilde{X} > x, \tilde{Y} > y]\, dC[F(x), G(y)] \tag{54}$$

$$= \iint\limits_{\mathbb{R}^2} \left\{1 - F(x) - G(y) + \tilde{C}[F(x), G(y)]\right\} dC[F(x), G(y)] \tag{55}$$

$$= \iint\limits_{[0,1]^2} \left\{1 - u - v + \tilde{C}(u,v)\right\} dC(u,v) \tag{56}$$

But since C is the joint distribution function of a vector $(U,V)^T$ of $U(0,1)$ random variables, $\mathbb{E}(U) = \mathbb{E}(V) = 1/2$, and hence

$$\mathbb{P}[X < \tilde{X}, Y < \tilde{Y}] = 1 - \frac{1}{2} - \frac{1}{2} + \iint\limits_{[0,1]^2} \tilde{C}(u,v)\, dC(u,v) = \iint\limits_{[0,1]^2} \tilde{C}(u,v)\, dC(u,v) \tag{57}$$

Thus

$$\mathbb{P}[(X - \tilde{X})(Y - \tilde{Y}) > 0] = 2 \iint\limits_{[0,1]^2} \tilde{C}(u,v)\, dC(u,v) \tag{58}$$

and the conclusion follows

$$Q = 2\mathbb{P}[(X - \tilde{X})(Y - \tilde{Y}) > 0] - 1 = 4 \iint\limits_{[0,1]^2} \tilde{C}(u,v)\, dC(u,v) - 1 \tag{59}$$

7.3 Kendall's tau and Spearman's rho

Definition. Kendall's tau for the random vector $(X,Y)^T$ is defined as

$$\tau(X,Y) = \mathbb{P}[(X - \tilde{X})(Y - \tilde{Y}) > 0] - \mathbb{P}[(X - \tilde{X})(Y - \tilde{Y}) < 0] \tag{60}$$

where $(\tilde{X}, \tilde{Y})^T$ is an independent copy of $(X,Y)^T$. Hence Kendall's tau for $(X,Y)^T$ is simply the probability of concordance minus the probability of discordance and since the copula of $(\tilde{X}, \tilde{Y})^T$ is the same of $(X,Y)^T$ is also simply equal to $Q(C,C)$:
Theorem. Let $(X,Y)^T$ be a vector of continuous random variables with copula C. Then Kendall's tau for $(X,Y)^T$ is given by

$$\tau(X,Y) = Q(C,C) = 4 \iint\limits_{[0,1]^2} C(u,v)\, dC(u,v) - 1 \tag{61}$$

Note that the integral above is the expected value of the random variable $C(U,V)$, where $U,V \sim U(0,1)$ with joint distribution function C, i.e. $\tau = 4\mathbb{E}(C(U,V)) - 1$.

Definition. Spearman's rho for the random vector $(X,Y)^T$ is defined as

$$\rho_S(X,Y) = 3(\mathbb{P}[(X - \tilde{X})(Y - Y') > 0] - \mathbb{P}[(X - \tilde{X})(Y - Y') < 0]) \tag{62}$$

where $(X,Y)^T$, $(\tilde{X}, \tilde{Y})^T$ and $(X', Y')^T$ are **independent** copies.
Theorem. Let $(X,Y)^T$ be a vector of continuous random variables with copula C. Then Spearman's rho for $(X,Y)^T$ is given by

$$\rho_S(X,Y) = 3Q(C,\Pi) = 12 \iint\limits_{[0,1]^2} uv\, dC(u,v) - 3 = 12 \iint\limits_{[0,1]^2} C(u,v)\, du\, dv - 3 \tag{63}$$

$$= \frac{\mathbb{E}(UV) - 1/4}{1/12} = \frac{Cov(U,V)}{\sqrt{Var(U)}\sqrt{Var(V)}} \tag{64}$$

$$= \rho[F(X), G(Y)] \tag{65}$$

8 Tail Dependence

For copulas without a simple closed form an alternative formula for λ_U is more useful.

Consider a pair of $U(0,1)$ random variables (U,V) with copula C. First note that

$$\mathbb{P}(V \le v \,|\, U = u) = \frac{\partial C(u,v)}{\partial u} \tag{66}$$

and

$$\mathbb{P}(V > v \,|\, U = u) = 1 - \frac{\partial C(u,v)}{\partial u} \tag{67}$$

and similarly when conditioning on V. Remember that, by definition

$$\frac{d\bar{C}(u,u)}{du} = \lim_{\Delta u \to 0} \frac{\bar{C}(u+\Delta u, u+\Delta u) - \bar{C}(u,u)}{\Delta u}$$

let's assume $\Delta u = 1 - u$

$$\frac{d\bar{C}(u,u)}{du} = \lim_{u \to 1} \frac{\bar{C}(1,1) - \bar{C}(u,u)}{1-u} = -\lim_{u \to 1} \frac{\bar{C}(u,u)}{1-u}$$

since $\bar{C}(1,1) = 0$

$$\begin{aligned}
\lambda_U &= \lim_{u \to 1} \frac{\bar{C}(u,u)}{1-u} = -\lim_{u \to 1} \frac{d\bar{C}(u,u)}{du} \\
&= \lim_{u \to 1} \left(-2 + \left.\frac{\partial C(s,t)}{\partial s}\right|_{s=t=u} + \left.\frac{\partial C(s,t)}{\partial t}\right|_{s=t=u} \right) \\
&= \lim_{u \to 1} \Big(\mathbb{P}(V > u \,|\, U = u) + \mathbb{P}(U > u \,|\, V = u) \Big)
\end{aligned} \tag{68}$$

Furthermore if C is an exchangeable copula, i.e. $C(u,v) = C(v,u)$, then the expression for λ_U simplifies to

$$\lambda_U = 2 \lim_{u \to 1} \mathbb{P}(V > u \,|\, U = u) \tag{69}$$

Example Let $(X,Y)^T$ have the bivariate standard normal distribution function with linear correlation coefficient ρ. That is $(X,Y)^T \sim C(\Phi(x), \Phi(y))$, where C is a member of the gaussian family. Since copulas in this family are exchangeable and because Φ is a distribution function with infinite right endpoint,

$$\lim_{u \to 1} \mathbb{P}\big(V > u \,|\, U = u\big) = \lim_{x \to \infty} \mathbb{P}\big(\Phi^{-1}(V) > x \,|\, \Phi^{-1}(U) = x\big) = \lim_{x \to \infty} \mathbb{P}\big(X > x \,|\, Y = x\big) \tag{70}$$

Using the well known fact that $(Y \,|\, X = x) \sim \mathcal{N}(\rho x, 1 - \rho^2)$ we obtain

$$\lambda_U = 2 \lim_{x \to \infty} \bar{\Phi}\Big((x - \rho x)/\sqrt{1-\rho^2}\Big) = 2 \lim_{x \to \infty} \bar{\Phi}\Big(x\sqrt{1-\rho}/\sqrt{1+\rho}\Big) \tag{71}$$

from which it follows that $\lambda_U = 0$ for $\rho < 1$, hence the gaussian copula does not have upper tail dependence.

If $(X_1, X_2)^T$ has a standard bivariate t-distribution with ν degrees of freedom and linear correlation matrix R, then $(X_2 | X_1 = x)$ is t-distributed with $\nu + 1$ degrees of freedom and

$$\mathbb{E}(X_2|X_1 = x) = R_{12} x, \quad Var(X_2|X_1 = x) = \left(\frac{\nu + x^2}{\nu + 1}\right)\Big(1 - R_{12}^2\Big)$$

This can be used to show that the t-copula has upper (and because of radial symmetry equal lower) tail dependence:

$$\begin{aligned}
\lambda_U &= 2 \lim_{x \to \infty} \mathbb{P}\big(X_2 > x | X_1 = x\big) \\
&= 2 \lim_{x \to \infty} \bar{t}_{\nu+1}\left(\left(\frac{\nu+1}{\nu+x^2}\right)^{1/2} \frac{x - R_{12}x}{\sqrt{1-R_{12}^2}} \right) \\
&= 2 \lim_{x \to \infty} \bar{t}_{\nu+1}\left(\left(\frac{\nu+1}{\nu/x^2+1}\right)^{1/2} \frac{\sqrt{1-R_{12}}}{\sqrt{1+R_{12}}} \right) \\
&= 2\bar{t}_{\nu+1}\left(\sqrt{\nu+1}\frac{\sqrt{1-R_{12}}}{\sqrt{1+R_{12}}} \right)
\end{aligned} \tag{72}$$

From this it is also seen that the coefficient of upper tail dependence is increasing in R_{12} and decreasing in ν. Furthermore, the coefficient of upper (lower) tail dependence tends to zero as the number of degrees of freedom tends to infinity for $R_{12} < 1$.

9 Exercises

Ex. 1 — Calculate the expected value and variance for a random variable with a log-normal distribution

Answer (Ex. 1) — By definition $Y = e^X$ where X is $N(\mu, \sigma)$. Therefore we can write

$$E[Y] = \alpha \int_{-\infty}^{+\infty} \exp(x) \exp\Big[-\frac{1}{2\sigma^2}(x-\mu)^2\Big]\, dx \tag{73}$$

being α a normalization factor. By computing the integrand we get

$$\begin{aligned}
\exp(x) \exp\Big[-\frac{1}{2\sigma^2}(x-\mu)^2\Big] &= \exp\Big[-\frac{1}{2\sigma^2}(x-\mu)^2 + x\Big] \\
&= \exp\Big[-\frac{1}{2\sigma^2}\Big((x-\mu)^2 - 2\sigma^2 x\Big)\Big] \\
&= \exp\Big[-\frac{1}{2\sigma^2}\Big(x^2 + \mu^2 - 2x\mu - 2\sigma^2 x\Big)\Big] \\
&= \exp\Big[-\frac{1}{2\sigma^2}\Big(x^2 + \mu^2 - 2x(\mu + \sigma^2)\Big)\Big]
\end{aligned} \tag{74}$$

Let's focus on the term in round brackets with the aim of reconstructing a binomial square

$$\begin{aligned}
x^2 + \mu^2 - 2x(\mu+\sigma^2) &= x^2 + \mu^2 - 2x(\mu+\sigma^2) + \sigma^4 - \sigma^4 + 2\mu\sigma^2 - 2\mu\sigma^2 \\
&= x^2 + (\mu+\sigma^2)^2 - 2x(\mu+\sigma^2) - \sigma^4 - 2\mu\sigma^2 \\
&= x^2 + (\mu+\sigma^2)^2 - 2x(\mu+\sigma^2) - \sigma^2(\sigma^2 + 2\mu)
\end{aligned} \tag{75}$$

Put $\nu = \mu + \sigma^2$, we can write the final result of (75) as

$$x^2 + \mu^2 - 2x(\mu+\sigma^2) = (x-\nu)^2 - \sigma^2(2\mu + \sigma^2)$$

So the integrand (74) becames

$$\begin{aligned}
\exp(x) \exp\Big[-\frac{1}{2\sigma^2}(x-\mu)^2\Big] &= \exp\Big[-\frac{1}{2\sigma^2}(x-\nu)^2 + \frac{1}{2\sigma^2}\sigma^2(2\mu+\sigma^2)\Big] \\
&= \exp\Big[-\frac{1}{2\sigma^2}(x-\nu)^2 + \frac{1}{2}(2\mu+\sigma^2)\Big] \\
&= \exp\Big[-\frac{1}{2\sigma^2}(x-\nu)^2 + \Big(\mu + \frac{\sigma^2}{2}\Big)\Big] \\
&= \exp\Big[-\frac{1}{2\sigma^2}(x-\nu)^2\Big] \exp\Big(\mu + \frac{\sigma^2}{2}\Big)
\end{aligned} \tag{76}$$

So the expected value can finally be written as

$$\begin{aligned}
E[Y] &= \alpha \int_{-\infty}^{+\infty} \exp(x) \exp\Big[-\frac{1}{2\sigma^2}(x-\mu)^2\Big]\, dx \\
&= \alpha \exp\Big(\mu + \frac{\sigma^2}{2}\Big) \int_{-\infty}^{+\infty} \exp\Big[-\frac{1}{2\sigma^2}(x-\nu)^2\Big]\, dx
\end{aligned} \tag{77}$$

But, by definition:

$$\alpha \int_{-\infty}^{+\infty} \exp\Big[-\frac{1}{2\sigma^2}(x-\nu)^2\Big]\, dx = 1$$

From this we have

$$E[Y] = \exp\Big(\mu + \frac{\sigma^2}{2}\Big) \tag{78}$$

For the calculation of the variance, let us first calculate the expected value of the square of Y. Taking into account that

$$Y^2 = \left(e^x\right)^2 = e^{2x}$$

we have

$$E[Y^2] = \alpha \int_{-\infty}^{+\infty} \exp(2x) \exp\Big[-\frac{1}{2\sigma^2}(x-\mu)^2\Big]\, dx \tag{79}$$

Therefore the calculation is completely analogous to the previous case except for a factor 2 that multiplies the exponent; so we can write

$$E[Y^2] = \exp\big(2\mu + 2\sigma^2\big) \tag{80}$$

Remember that

$$var(x) = E\big(x^2\big) - \big[E(x)\big]^2$$

We get

$$\begin{aligned} var(x) &= \exp\big(2\mu + 2\sigma^2\big) - \exp\Big[2\Big(\mu + \frac{\sigma^2}{2}\Big)\Big] \\ &= \exp\big(2\mu + \sigma^2\big)\exp\big(\sigma^2\big) - \exp\big(2\mu + \sigma^2\big) \\ &= \exp\big(2\mu + \sigma^2\big)\Big[\exp\big(\sigma^2\big) - 1\Big] \end{aligned} \tag{81}$$

From this we obtain also

$$\sigma(x) = \exp\left(\mu + \frac{\sigma^2}{2}\right)\sqrt{\exp\big(\sigma^2\big) - 1} \tag{82}$$

Ex. 2 — Let X_1 and X_2 have a bivariate normal distribution with joint probability density function

$$\begin{aligned} f(x_1, x_2) = {} & \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}} \\ & \times \exp\Big\{-\frac{1}{2(1-\rho^2)}\Big[\Big(\frac{x_1-\mu_1}{\sigma_1}\Big)^2 + \Big(\frac{x_2-\mu_2}{\sigma_2}\Big)^2 \\ & - 2\rho\Big(\frac{x_1-\mu_1}{\sigma_1}\Big)\Big(\frac{x_2-\mu_2}{\sigma_2}\Big)\Big]\Big\} \end{aligned} \tag{83}$$

Now compute $Cov(Y_1, Y_2)$ where $Y_1 = e^{x_1}$ and $Y_2 = e^{x_2}$

Answer (Ex. 2) — We can write

$$Cov(Y_1, Y_2) = \mathbb{E}(Y_1Y_2) - \mathbb{E}(Y_1)\mathbb{E}(Y_2) \tag{84}$$

First of all note that since $Y_1Y_2 = \exp(X_1 + X_2)$, then $\log(Y_1Y_2)$ has a $N(\mu_1+\mu_2, \sigma_1^2+\sigma_2^2+2\rho\sigma_1\sigma_2)$ distribution therefor, using the result of the previous exercise, we can write:

$$\mathbb{E}(Y_1Y_2) = exp\left[(\mu_1+\mu_2) + \frac{1}{2}(\sigma_1^2 + \sigma_2^2 + 2\rho\sigma_1\sigma_2)\right] \tag{85}$$

Therefore

$$\begin{aligned} Cov(Y_1, Y_2) &= \exp\left[(\mu_1+\mu_2) + \frac{1}{2}(\sigma_1^2 + \sigma_2^2 + 2\rho\sigma_1\sigma_2)\right] - exp\left[\mu_1 + \frac{\sigma_1^2}{2} + \mu_2 + \frac{\sigma_2^2}{2}\right] \\ &= \exp\left[\mu_1 + \mu_2 + \frac{1}{2}(\sigma_1^2 + \sigma_2^2)\right]\left\{\exp\big(\rho\sigma_1\sigma_2\big) - 1\right\} \end{aligned} \tag{86}$$

Therefore the linear correlation coefficient of Y_1 and Y_2 is

$$\begin{aligned}\rho_{Y_1,Y_2} &= \frac{Cov(Y_1,Y_2)}{\sigma(Y_1)\sigma(Y_2)} \\ &= \frac{\exp\left[\mu_1+\mu_2+\frac{1}{2}(\sigma_1^2+\sigma_2^2)\right]\left\{\exp\big(\rho\sigma_1\sigma_2\big)-1\right\}}{\exp\left(\mu_1+\frac{\sigma_1^2}{2}\right)\sqrt{\exp\big(\sigma_1^2\big)-1}\,\exp\left(\mu_2+\frac{\sigma_2^2}{2}\right)\sqrt{\exp\big(\sigma_2^2\big)-1}} \\ &= \frac{\exp\big(\rho\sigma_1\sigma_2\big)-1}{\sqrt{\exp\big(\sigma_1^2\big)-1}\sqrt{\exp\big(\sigma_2^2\big)-1}}\end{aligned} \tag{87}$$

Ex. 3 — Compute the following expression

$$\int_0^1\int_0^1 C(u,v)\frac{\partial^2 C(u,v)}{\partial u\partial v}\,du\,dv \tag{88}$$

Answer (Ex. 3) — Evaluate the inner integral by parts

$$\int_0^1 C(u,v)\frac{\partial^2}{\partial u\partial v}C(u,v)\,du \tag{89}$$

$$= C(u,v)\frac{\partial}{\partial v}C(u,v)\Big|_{u=0}^{u=1} - \int_0^1 \frac{\partial}{\partial u}C(u,v)\frac{\partial}{\partial v}C(u,v)\,du \tag{90}$$

$$= v - \int_0^1 \frac{\partial}{\partial u}C(u,v)\frac{\partial}{\partial v}C(u,v)\,du \tag{91}$$

$$\int_0^1\int_0^1 C(u,v)\frac{\partial^2}{\partial u\partial v}C(u,v)\,dudv \tag{92}$$

$$= \frac{1}{2} - \int_0^1\int_0^1 \frac{\partial}{\partial u}C(u,v)\frac{\partial}{\partial v}C(u,v)\,dudv. \tag{93}$$