DENSITY OF THE SUM OF DEPENDENT RANDOM VARIABLES VIA COPULA FUNCTIONS

Dushatov N.T.
Almalyk state technical institute
n_dushatov@rambler.ru

Abstract: In this paper, we investigate the distribution of the sum of two dependent random variables whose marginal densities are known. The dependence structure between the variables is modeled using a copula function. Based on Sklar's theorem, we derive an explicit integral representation for the probability density function of the sum. The obtained result generalizes the classical convolution formula for independent random variables and provides a flexible framework for modeling dependence in applied probability and statistics [1], [2].

1. Introduction

The problem of determining the distribution of the sum of random variables is fundamental in probability theory. When random variables are independent, the distribution of their sum is obtained via the classical convolution formula. However, in many real-world applications, such as finance, insurance, and reliability theory, random variables exhibit dependence.

Copula theory provides a powerful tool to separate marginal behavior from dependence structure. By employing copulas, one can construct multivariate distributions with prescribed marginal distributions and flexible dependence patterns. The objective of this paper is to derive the probability density function of the sum of two dependent random variables using copula functions. [1], [3]

2. Preliminaries and Definitions

Let X and Y be continuous random variables with marginal distribution functions F_X and F_Y, and corresponding density functions f_X and f_Y.

A copula is a function $C:[0,1]^2 \to [0,1]$ that couples univariate distribution functions to form a joint distribution function.

Theorem 2.1 (Sklar's Theorem).[3-4] For any joint distribution function $F_{X,Y}$ with continuous marginals F_X and F_Y, there exists a unique copula C such that

$$F_{X,Y}(x,y) = C\left(F_X(x), F_Y(y)\right).$$

3. Joint Density Representation via Copulas

Assume that the copula $C(u,v)$ is twice continuously differentiable. Define the copula density as

$$c(u,v) = \frac{\partial^2 C(u,v)}{\partial u \partial v}.$$

Theorem 3.1[1,5]. The joint density function of X and Y is given by

$$f_{X,Y}(x,y) = c\left(F_X(x), F_Y(y)\right) f_X(x) f_Y(y).$$

Proof.

By Sklar's theorem,

$$F_{X,Y}(x,y) = C\left(F_X(x), F_Y(y)\right).$$

Differentiating with respect to x and y and applying the chain rule yields the desired result. ■

4. Density of the Sum of Dependent Random Variables

Let $Z = X + Y$

Theorem 4.1 [2,6] The probability density function of Z is given by

$$f_Z(z) = \int_{-\infty}^{\infty} c\left(F_X(x), F_Y(z-x)\right) f_X(x) f_Y(z-x) dx.$$

Proof.

By definition,

$$f_Z(z) = \int_{-\infty}^{\infty} f_{XY}(x, z-x) dx.$$

Substituting the joint density from Theorem 3.1 completes the proof. ■

4A. Example: Gaussian Copula Case

Consider a Gaussian copula with parameter $\rho \in (-1,1)$. Let $\Phi(g)$ and $\varphi(g)$ denote the standard normal distribution function and density, respectively. The Gaussian copula density can be written as

$$c_\rho(u,v) = \frac{1}{\sqrt{1-\rho^2}} \exp\left\{-\frac{\rho^2(z_1^2 + z_2^2) - 2\rho z_1 z_2}{2(1-\rho^2)}\right\},$$

where $z_1 = \Phi^{-1}(u)$, $z_2 = \Phi^{-1}(v)$. Substituting $u = F_X(x), v = F_Y(z-x)$ into Theorem 4.1 yields

$$f_Z(z) = \int_{-\infty}^{\infty} c_\rho\left(F_X(x), F_Y(z-x)\right) f_X(x) f_Y(z-x) dx.$$

In particular, for standard normal marginal $X \sim N(0,1)$, $Y \sim N(0,1)$, we have $F_X(x) = \Phi(x)$, $F_Y(y) = \Phi(y)$, hence

$$z_1 = \Phi^{-1}\left(\Phi(x)\right) = x, \; z_2 = \Phi^{-1}\left(\Phi(z-x)\right) = z - x,$$

and therefore an explicit integrand is obtained:

$$f_Z(z) = \int_{-\infty}^{\infty} \frac{1}{\sqrt{1-\rho^2}} \exp\left\{\frac{2\rho x(z-x) - \rho^2(x^2 + (z-x)^2)}{2(1-\rho^2)}\right\} \varphi(x)\varphi(z-x) dx.$$

Although a closed form is generally not available, the integral representation is fully explicit and can be efficiently evaluated numerically for given ρ, thus enabling a direct study of the dependence impact on the distribution of Z.

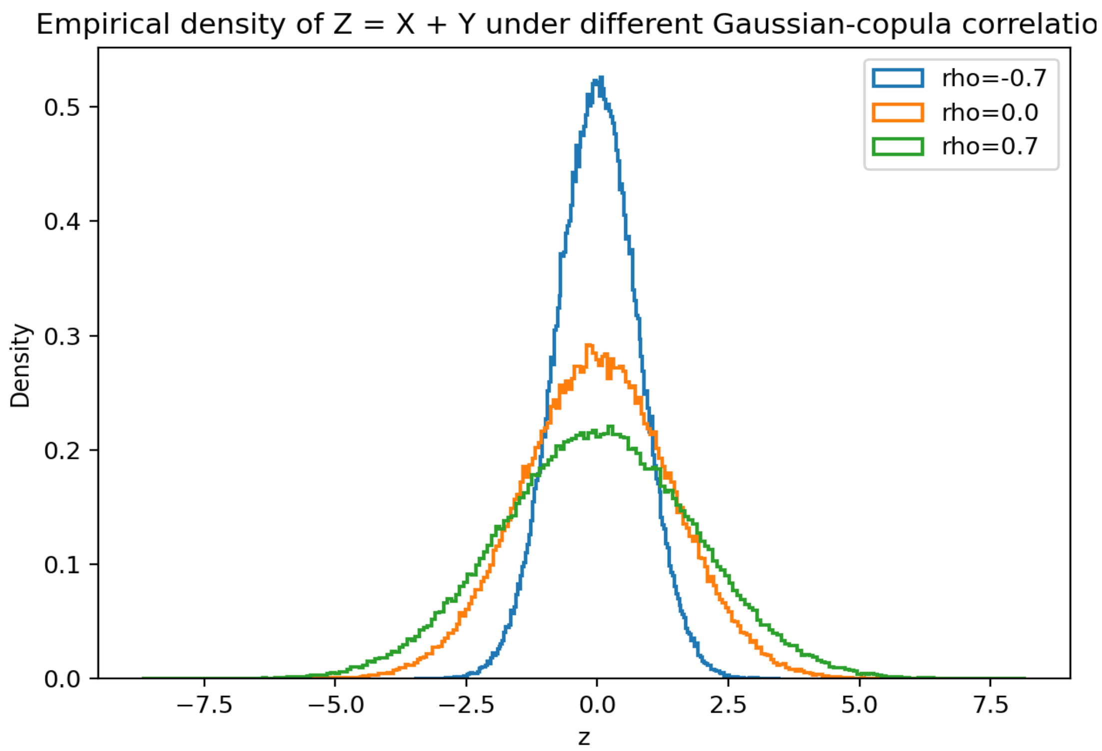
4B. Numerical Simulation Study

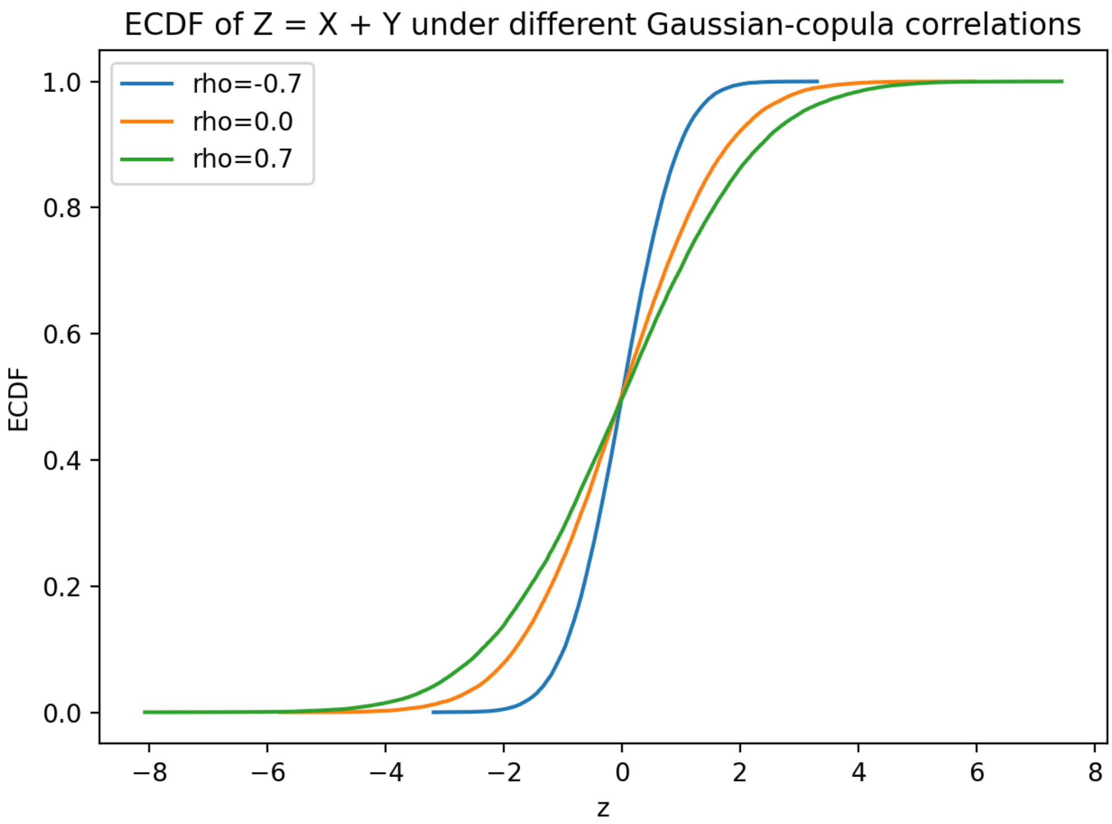
To validate the theoretical findings and illustrate the effect of dependence, we performed a Monte Carlo study under standard normal marginals. For each $\rho \in \{-0.7; 0; 0.7\}$, we generated $N = 200000$ samples (X_i, Y_i) from a bivariate normal distribution with correlation ρ, which corresponds to a Gaussian copula with normal marginals. We then computed $Z_i = X_i + Y_i$ and compared empirical summaries and distributional shapes.

Table 1 reports the empirical mean, variance, skewness, and excess kurtosis of Z. As expected, the mean remains close to zero, while the variance increases with positive dependence and decreases with negative dependence. This behavior is also visible in **Figure 1** (empirical densities) and **Figure 2** (ECDFs): positive ρ produces a wider distribution (more mass in the tails), whereas negative ρ yields concentration around the center.

Table 1. Monte Carlo summary for $Z = X + Y$ under Gaussian copula

Copula parameter (ρ)	Mean (Z)	Variance (Z)	Skewness (Z)	Excess Kurtosis (Z)
-0.7	0.000355	0.597737	-0.002216	0.009778
0.0	-0.002105	2.008798	-0.001089	-0.007034
0.7	0.005487	3.396974	-0.000472	-0.014695





4C. Extension to the Multivariate Case.

The approach extends naturally to the sum of n dependent continuous random variables. Let $X_1, X_2, ..., X_n$ have marginal densities f_i and distribution functions F_i, and assume the joint distribution is characterized by an $n-$dimensional copula C with density

$$c(u_1, u_2, ..., u_n) = \frac{\partial^n}{\partial u_1 \cdots \partial u_n} C(u_1, u_2, ..., u_n).$$

Then the joint density is

$$f_{X_1,...,X_n}(x_1,...,x_n) = c(F_1(x_1),...,F_n(x_n)) \prod_{i=1}^{n} f_i(x_i).$$

For the sum $Z = \sum_{i=1}^{n} X_i$, an integral representation of the density is obtained by integrating out $n-1$ variables:

$$f_Z(z) = \int_{R^{n-1}} c\left(F_1(x_1),...,F_{n-1}(x_{n-1}), F_n\left(z - \sum_{i=1}^{n-1} x_i\right)\right)\left(\prod_{i=1}^{n-1} f_i(x_i)\right) f_n\left(z - \sum_{i=1}^{n-1} x_i\right) dx_1 \cdots dx_{n-1}.$$

This representation is particularly useful in risk aggregation and portfolio modeling, where sums of multiple dependent components arise naturally.

5. Discussion

The derived formula demonstrates that the classical convolution is a special case corresponding to the independence copula. Different copulas lead to different dependence effects on the distribution of the sum. This framework is applicable in finance, insurance risk modeling, and reliability analysis [1], [6].

6. Conclusion

In this paper, an explicit integral representation for the density of the sum of two dependent random variables was obtained using copula theory. The result extends the classical convolution approach and provides a flexible tool for dependence modeling in probability and statistics [1], [2].

References:

1. Nelsen, R. B. *An Introduction to Copulas*. Springer.
2. Joe, H. *Dependence Modeling with Copulas*. Chapman & Hall. 2014. CRC Press.
3. Sklar, A. Fonctions de répartition à n dimensions et leurs marges, *Publ. Inst. Statist. Univ. Paris*, 1959.
4. Schweizer, B., Sklar, A. *Probabilistic Metric Spaces*. Dover.
5. Joe, H. (1997). Multivariate Models and Dependence Concepts. Chapman & Hall, London.
6. Durante, F., Sempi, C. *Principles of Copula Theory*. CRC Press.