The sum of dependent normal variables may be not normal

A. Novosyolov*

August 14, 2006

Abstract

This brief note provides calculation of distributions in the example of dependent normal variables with non-normal sum. The example belongs to Glyn Holton [2] and intended to break the illusion of normality of linear combinations of normal variables. Comments are solicited.

Contents

*Institute of computational modeling SB RAS, 660036, Krasnoyarsk, Academgorodok, e-mail: anov@inbox.ru

1 Introduction

Let a random vector $(X_1, \ldots, X_n)$ possess a joint normal distribution, i.e. its probability density function (PDF) has the form

$$f(x) = \frac{|S|^{1/2}}{(2\pi)^{n/2}} \exp\left(-\frac{1}{2}(x-\mu)^T S(x-\mu)\right), \quad x \in \mathbf{R}^n, \tag{1}$$

where $\mu \in \mathbf{R}^n$ is an n-dimensional vector (vector of distribution means), S is a positive definite $n \times n$ matrix (its inverse S^{-1} is a covariance matrix), $|S|$ stands for the determinant of S, and upper index T denotes transposition. Such a vector is often called simply normal vector. In the special case $n = 1$ one obtains a univariate normal distribution with PDF

$$f(x) = \frac{1}{\sqrt{2\pi\sigma}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right), \quad x \in \mathbf{R} \tag{2}$$

with $\mu \in \mathbf{R}$ and $\sigma^2 > 0$ denoting the distribution mean and variance. The corresponding variable is called simply (univariate) normal.

If $\mu = 0$ and $\sigma = 1$, the distribution described by the PDF (2) is called standard normal. Similarly, if $\mu = (0, \ldots, 0)^T$, and S is a unit matrix, the distribution with PDF (1) is called (multivariate) standard normal. Note that the components of the latter are independent.

The following theorem is well known, see eg. [1].

Theorem 1 *A random vector $(X_1, \ldots, X_n)$ has a joint normal distribution if and only if* each *linear combination*

$$a_1 X_1 + \cdots + a_n X_n \tag{3}$$

has a (univariate) normal distribution, $a_1, \ldots, a_n \in \mathbf{R}$.

Theorem 1 in particular means that the components $X_1, \ldots, X_n$ of the normal random vector are themselves normal variables. Indeed, X_1 is a linear combination of the form $1 \cdot X_1 + 0 \cdot X_2 + \cdots + 0 \cdot X_n$; similar expressions also hold for other components. The theorem also means that the sum $X_1 + \cdots + X_n$ of the components of a normal random vector is itself normal.

This fact often provokes an illusion that a sum of normal variables should be normal. Sometimes this may be true, eg. in case of independent summands (in this case the joint PDF equals the product of marginal normal PDFs of the form (2) and has the form (1) with diagonal matrix S). This is also true in case of any joint normal distribution, as theorem 1 states.

However in general normality of marginal variables $X_1, \ldots, X_n$ does not imply normality of their joint distribution and thus does not imply normality of their sum. An example of the sort was described in [2, pp. 138,139], see also [3]. The example has not been supplied with proofs. This causes doubts in its validity, eg. [4]. The present note is devoted to proving key statements of the example.

Section 2 contains the statement of the problem from [2, 3]. In section 3 we prove one general theorem, and in section 4 we supply all necessary proofs. Section 5 discusses the topic using the concept of copula function.

2 Problem statement

Recall that our goal is to construct an example of random vector (X_1, X_2) such that components X_1 and X_2 are univariate normal, but the distribution of $X_1 + X_2$ is not normal.

In this section we describe the example from [2], [3]. The starting point in this example is a pair of independent normal standard random variables X_1 and Z. In other words the random vector (X_1, Z) possesses the standard bivariate normal distribution.

Now consider the random variable

$$X_2 = \begin{cases} |Z|, & X_1 \geq 0 \\ -|Z|, & X_1 < 0 \end{cases} \tag{4}$$

Its distribution is standard normal, yet the joint distribution of (X_1, X_2) is not normal. The latter may be concluded from the obvious relation $\mathbf{P}(X_1 X_2 \geq 0) = 1$, which is impossible in case of normality. Moreover, the distribution of $X_1 + X_2$ is not normal.

The two statements in the above paragraph may be not obvious and may cause doubt:

- The distribution of X_2 is univariate normal;
- The distribution of $X_1 + X_2$ is not a univariate normal.

The rest of the note is devoted to proving these two statements.

3 A general theorem

To simplify things we would abstract from unnecessary details and consider more general statement in this section.

Theorem 2 *Let X_1, Z be a pair of independent random variables such that the cumulative distribution function (CDF) F of Z is continuous:*

$$F(x) = \mathbf{P}(Z \leq x) = \mathbf{P}(Z < x), \quad -\infty < x < \infty \tag{5}$$

and symmetric with respect to 0:

$$F(x) = 1 - F(-x), \quad -\infty < x < \infty. \tag{6}$$

Next, let

$$\mathbf{P}(X_1 \geq 0) = \mathbf{P}(X_1 < 0) = \frac{1}{2}, \tag{7}$$

and the variable X_2 be defined by (4). Then the CDF of X_2 is equal to that of Z, i.e., F. In other words,

$$\mathbf{P}(X_2 \leq x) = F(x), \quad -\infty < x < \infty. \tag{8}$$

Proof. We will prove (8) using (4) — (7). First, by the law of total probability, using (7), we obtain

$$\begin{aligned}\mathbf{P}(X_2 \leq x) &= \mathbf{P}(X_2 \leq x \,|\, X_1 \geq 0)\mathbf{P}(X_1 \geq 0) + \mathbf{P}(X_2 \leq x \,|\, X_1 < 0)\mathbf{P}(X_1 < 0) \\ &= \frac{1}{2}\left[\mathbf{P}(X_2 \leq x \,|\, X_1 \geq 0) + \mathbf{P}(X_2 \leq x \,|\, X_1 < 0)\right]. \end{aligned} \tag{9}$$

Now we calculate the terms in brackets separately. From (4) it is clear that $X_1 \geq 0$ implies $X_2 = |Z|$, so

$$\mathbf{P}(X_2 \leq x \,|\, X_1 \geq 0) = \mathbf{P}(|Z| \leq x \,|\, X_1 \geq 0).$$

Since X_1 and Z are independent, the last conditional probability coincides with the unconditional one, so

$$\mathbf{P}(X_2 \leq x \,|\, X_1 \geq 0) = \mathbf{P}(|Z| \leq x). \tag{10}$$

In case $x < 0$ clearly $\mathbf{P}(|Z| \leq x) = 0$, so consider the case $x \geq 0$. Using continuity (5) of F we have

$$\mathbf{P}(|Z| \leq x) = \mathbf{P}(-x \leq Z \leq x) = \mathbf{P}(-x < Z \leq x) = F(x) - F(-x).$$

Now symmetry (6) implies

$$\mathbf{P}(|Z| \leq x) = F(x) - 1 + F(x) = 2F(x) - 1.$$

Thus

$$\mathbf{P}(X_2 \leq x \,|\, X_1 \geq 0) = \begin{cases} 0, & x < 0; \\ 2F(x) - 1, & x \geq 0. \end{cases} \tag{11}$$

Now consider the second conditional probability in (9). Using the definition (4) and independence of X_1 and Z, we have

$$\mathbf{P}(X_2 \leq x \,|\, X_1 < 0) = \mathbf{P}(-|Z| \leq x \,|\, X_1 < 0) = \mathbf{P}(-|Z| \leq x).$$

Clearly $\mathbf{P}(-|Z| \le x) = 1$ for $x \ge 0$, so consider the case $x < 0$. Since $-x > 0$ and $(|Z| \ge a) = (Z \le -a) \cup (Z \ge a)$ for $a > 0$, we have

$$\mathbf{P}(-|Z| \le x) = \mathbf{P}(|Z| \ge -x) = \mathbf{P}(Z \le x) + \mathbf{P}(Z \ge -x).$$

Now, using continuity and symmetry of F again, one obtains

$$\mathbf{P}(-|Z| \le x) = \mathbf{P}(Z \le x) + \mathbf{P}(Z > -x) = F(x) + 1 - F(-x) = 2F(x).$$

Finally,

$$\mathbf{P}(X_2 \le x \,|\, X_1 < 0) = \begin{cases} 2F(x), & x < 0; \\ 1, & x \ge 0. \end{cases} \tag{12}$$

Now substitution of (11) and (12) into (9) gives $\mathbf{P}(X_2 \le x) = F(x)$ for all $-\infty < x < \infty$ as required. $\diamond$

4 Completing the example

4.1 X_2 is normal

The statement follows immediately from the theorem 2. Indeed, CDF of the standard normal variable Z possesses the properties of continuity and symmetry (5), (6), and the standard normal variable X_1 clearly satisfies (7). So by theorem 2 the random variable X_2 defined by (4) has the same distribution as Z, i.e. standard normal.

4.2 $X_1 + X_2$ is not normal

Proof of this fact is a bit more complicated, so we first demonstrate it by Monte Carlo simulations. Figure 1 depicts the histogram of the distribution of $X_1 + X_2$ obtained by 200,000 Monte Carlo trials, and the normal PDF with the same mean and variance as of $X_1 + X_2$. Clearly the distribution of $X_1 + X_2$ does not resemble normal distribution.

Now let's give a formal proof. Denote φ the PDF of the univariate standard normal distribution:

$$\varphi(x) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{x^2}{2}\right), \quad -\infty < x < \infty$$

and denote

$$\Phi(x) = \int_{-\infty}^{x} \varphi(t)\, dt, \quad -\infty < x < \infty$$

its CDF. Next, denote g the PDF of the bivariate standard normal distribution

$$g(x, y) = \frac{1}{2\pi} \exp\left(-\frac{x^2 + y^2}{2}\right).$$

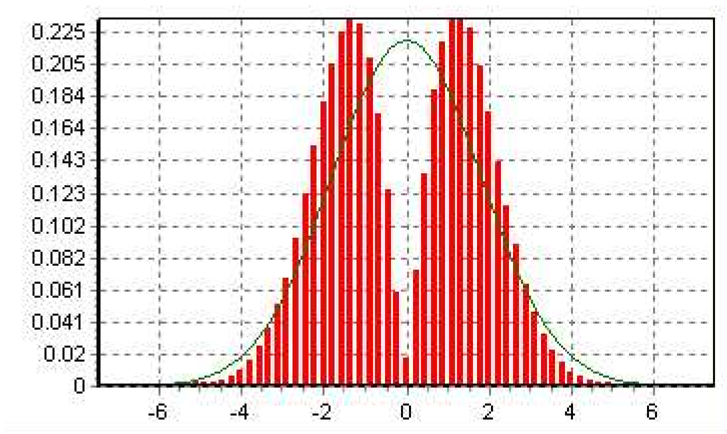


Figure 1: Histogram for the distribution of $X_1 + X_2$ by 200,000 Monte Carlo trials (red); normal density (green)

In particular, this is the PDF of the pair (X_1, Z) defined in section 2.

From reflection nature of the transform (4) it is clear, that the PDF f_{X_1,X_2} of (X_1, X_2) is twice as large as that of (X_1, Z) in areas $xy > 0$ and equals to 0 in areas[1] $xy < 0$, see figure 2.

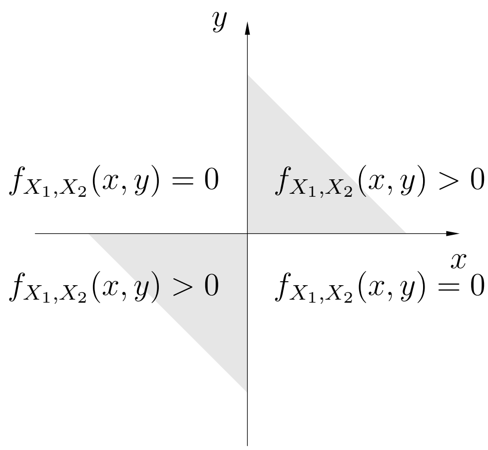


Figure 2: Areas of positive values of the PDF f_{X_1,X_2}

[1]The PDF f_{X_1,X_2} is discontinuous and undefined on both axes $x = 0$ and $y = 0$, so we may define it as appropriate there, since this does not affect the distribution of (X_1, X_2)

Formally:

$$f_{X_1,X_2}(x,y) = \begin{cases} \frac{1}{\pi}\exp\left(-\frac{x^2+y^2}{2}\right), & xy > 0; \\ 0, & \text{otherwise.} \end{cases} \tag{13}$$

Now, the CDF of $X_1 + X_2$ may be calculated by

$$F_{X_1+X_2}(z) = \mathbf{P}(X_1 + X_2 \le z) = \int_{x+y\le z} f_{X_1,X_2}(x,y)\,dx\,dy, \quad -\infty < z < \infty.$$

Since f_{X_1,X_2} is symmetric with respect to the origin, it is clear that the distribution of X_1+X_2 is symmetric with respect to 0, so $F_{X_1+X_2}(z) = 1 - F_{X_1+X_2}(-z)$, and it suffices to calculate it for $z \ge 0$. Moreover, it is clear that $F_{X_1+X_2}$ if continuous, so $F_{X_1+X_2}(0) = 1/2$.

Now, for $z > 0$, we have

$$F_{X_1+X_2}(z) = \frac{1}{2} + \int_{x\ge 0, y\ge 0, x+y\le z} f_{X_1,X_2}(x,y)\,dx\,dy. \tag{14}$$

The integration area is shown in figure 3. To continue we need the following lemmas.

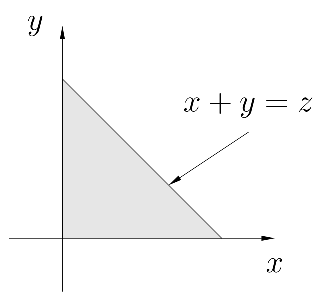


Figure 3: Integration area in (14)

Lemma 1 *For* $z \ge 0$

$$\frac{1}{2\pi}\int_0^z \int_0^{z-x} \exp\left(-\frac{x^2+y^2}{2}\right) dy\,dx = \int_0^z \varphi(x)\Phi(z-x)\,dx - \frac{1}{2}\Phi(z) + \frac{1}{4}\,.$$

Proof

$$\begin{aligned} &\frac{1}{2\pi}\int_0^z \int_0^{z-x} \exp\left(-\frac{x^2+y^2}{2}\right) dy\,dx \\ &= \frac{1}{\sqrt{2\pi}}\int_0^z \exp\left(-\frac{x^2}{2}\right)\left[\frac{1}{\sqrt{2\pi}}\int_0^{z-x} \exp\left(-\frac{y^2}{2}\right)\,dy\right] dx \end{aligned}$$

$$= \frac{1}{\sqrt{2\pi}} \int_0^z \exp\left(-\frac{x^2}{2}\right) \left[\Phi(z-x) - \frac{1}{2}\right] dx$$

$$= \int_0^z \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{x^2}{2}\right) \Phi(z-x)\, dx - \frac{1}{2}\left(\Phi(z) - \frac{1}{2}\right)$$

$$= \int_0^z \varphi(x)\Phi(z-x)\, dx - \frac{1}{2}\Phi(z) + \frac{1}{4}$$

as required. ⋄

Lemma 2 *For $z \geq 0$*

$$\int_0^z \varphi(x)\varphi(z-x)\, dx = \frac{1}{\sqrt{2}}\varphi\left(\frac{z}{\sqrt{2}}\right)\left[2\Phi\left(\frac{z}{\sqrt{2}}\right) - 1\right].$$

Proof We have

$$\int_0^z \varphi(x)\varphi(z-x)\, dx$$

$$= \frac{1}{2\pi}\int_0^z \exp\left(-\frac{x^2}{2}\right)\exp\left(-\frac{(z-x)^2}{2}\right) dx$$

$$= \frac{1}{2\pi}\int_0^z \exp\left(-x^2 + xz - \frac{z^2}{4} - \frac{z^2}{4}\right) dx$$

$$= \frac{1}{2\pi}\int_0^z \exp\left(-\left(x - \frac{z}{2}\right)^2\right)\exp\left(-\frac{z^2}{4}\right) dx$$

$$= \frac{1}{\sqrt{2\pi}}\exp\left(-\frac{z^2}{4}\right)\int_0^z \frac{1}{\sqrt{2\pi}}\exp\left(-\left(x - \frac{z}{2}\right)^2\right) dx$$

Now applying the change of variables $y = \sqrt{2}x - z/\sqrt{2}$ leads to

$$\int_0^z \varphi(x)\varphi(z-x)\, dx$$

$$= \frac{1}{\sqrt{2}}\frac{1}{\sqrt{2\pi}}\exp\left(-\frac{z^2}{4}\right)\int_{-z/\sqrt{2}}^{z/\sqrt{2}} \frac{1}{\sqrt{2\pi}}\exp\left(-\frac{y^2}{2}\right) dy$$

$$= \frac{1}{\sqrt{2}}\varphi\left(\frac{z}{\sqrt{2}}\right)\left[\Phi\left(\frac{z}{\sqrt{2}}\right) - \Phi\left(-\frac{z}{\sqrt{2}}\right)\right]$$

$$= \frac{1}{\sqrt{2}}\varphi\left(\frac{z}{\sqrt{2}}\right)\left[2\Phi\left(\frac{z}{\sqrt{2}}\right) - 1\right]$$

as required. ⋄

Let's calculate the integral in (14). Applying lemma 1, we obtain

$$\begin{aligned}
&\int_{x\geq 0, y\geq 0, x+y\leq z} f_{X_1,X_2}(x,y)\,dx\,dy \\
&= \frac{1}{\pi}\int_0^z \int_0^{z-x} \exp\left(-\frac{x^2+y^2}{2}\right) dy\,dx \\
&= 2\int_0^z \varphi(x)\Phi(z-x)\,dx - \Phi(z) + \frac{1}{2}\,.
\end{aligned}$$

Now (14) implies

$$F_{X_1+X_2}(z) = 1 - \Phi(z) + 2\int_0^z \varphi(x)\Phi(z-x)\,dx. \tag{15}$$

Differentiating this with respect to z provides the PDF of $X_1 + X_2$:

$$f_{X_1+X_2}(z) = -\varphi(z) + 2\varphi(z)\Phi(0) + 2\int_0^z \varphi(x)\varphi(z-x)\,dx = 2\int_0^z \varphi(x)\varphi(z-x)\,dx$$

Now lemma 2 gives

$$f_{X_1+X_2}(z) = \sqrt{2}\,\varphi\left(\frac{z}{\sqrt{2}}\right)\left[2\Phi\left(\frac{z}{\sqrt{2}}\right) - 1\right]. \tag{16}$$

We calculated $f_{X_1+X_2}(z)$ for $z \geq 0$. Due to symmetry of the distribution of $X_1 + X_1$ we have $f_{X_1+X_2}(z) = f_{X_1+X_2}(-z)$. Finalize the result in the following

Proposition 1 *Let (X_1, Z) be standard normal bivariate vector and X_2 be defined by (4). Then the PDF of $X_1 + X_2$ is expressed by*

$$f_{X_1+X_2}(z) = \sqrt{2}\,\varphi\left(\frac{z}{\sqrt{2}}\right)\left[2\Phi\left(\frac{|z|}{\sqrt{2}}\right) - 1\right].$$

Figure 4 reproduces the figure 1 with added graph of $f_{X_1+X_2}$. Note that in particular $f_{X_1+X_2}(0) = 0$ and

$$f_{X_1+X_2}(z) \sim \sqrt{2}\,\varphi\left(\frac{z}{\sqrt{2}}\right) \quad \text{as} \quad |z| \to \infty.$$

5 Description in terms of copulas

Let's briefly discuss what can be said on the topic from the copulas perspective. An introduction to the concept may be found in [5]. Recall that copula function C is a mapping from $[0,1]^n$ to $[0,1]$ and may be defined as a CDF on $[0,1]^n$ with uniform marginals.

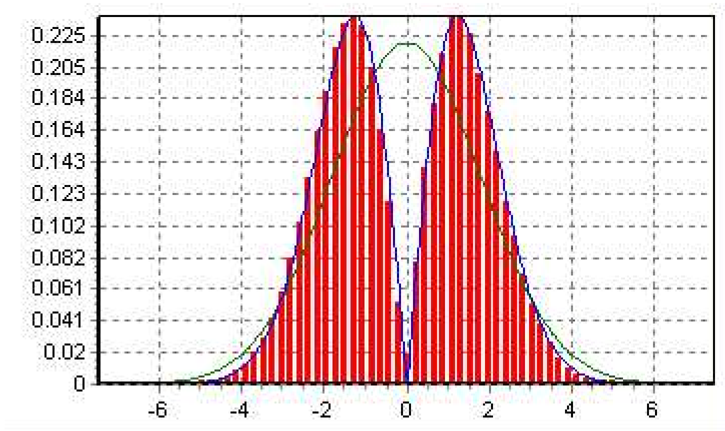


Figure 4: Histogram for the distribution of $X_1 + X_2$ by 200,000 Monte Carlo trials (red); normal PDF (green) and $f_{X_1+X_2}$ (blue)

5.1 Copulas in general

If $(X_1, \ldots, X_n)$ is a random vector with joint CDF $F_{X_1,\ldots,X_n}$ and continuous marginal CDFs F_{X_i}, $i = 1, \ldots, n$, then the copula function of the distribution of the random vector may be calculated by

$$C(u_1, \ldots, u_n) = F_{X_1,\ldots,X_n}\left(F_{X_1}^{-1}(u_1), \ldots, F_{X_n}^{-1}(u_n)\right),$$

where $F_{X_i}^{-1}$, $i = 1, \ldots, n$ stand for inverse CDFs. The inverse statement is Sklar's theorem

$$F_{X_1,\ldots,X_n}(x_1, \ldots, x_n) = C(F_1(x_1), \ldots, F_n(x_n)), \tag{17}$$

which states that the copula function C in the representation is unique in case of continuous marginals. In case of discontinuous marginals the representation remains valid, but copula function is not unique.

If the derivative

$$c(U_1, \ldots, u_n) = \frac{\partial^n C(u_1, \ldots, u_n)}{\partial u_1 \cdots \partial u_n}$$

exists, it represents the PDF of the copula C.

A very special class of copulas is formed by normal copulas; they have the form

$$C_\Sigma(u_1, \ldots, u_n) = \Phi_\Sigma\left(\Phi^{-1}(u_1), \ldots, \Phi^{-1}(u_n)\right),$$

where Φ_Σ is the CDF of the multivariate normal distribution with zero means, unit variances and correlation matrix Σ, and Φ^{-1} denotes the inverse univariate standard normal CDF. Normal copulas C_Σ possess PDFs c_Σ.

There is a whole lot of non-normal copulas widely used in finance, see eg. [6].

The following fact demonstrates an interesting property of copulas, which may be called invariance with respect to strictly monotone transforms. Let $(X_1, \dots, X_n)$ be any random vector with continuous marginal CDFs (so the copula in (17) is unique). Let the functions $g_1, \dots, g_n$ be strictly increasing. Denote $Y_1 = g_1(X_1), \dots, Y_n = g_n(X_n)$. Then the random vector $(Y_1, \dots, Y_n)$ possesses the same copula function as that of $(X_1, \dots, X_n)$. In particular, shifting and scaling components does not change the copula. That is why normal copula depends only on correlation matrix of the underlying joint normal distribution.

5.2 Copulas in the example

Now consider what the copula concept may bring to our topic of interest. The following statements are valid.

1. Combining normal copula $C = C_\Sigma$ with normal marginals (arbitrary shifted and scaled) in (17) provides joint normal CDF;
2. Combining normal copula with non-normal marginals in (17) provides non-normal CDF;
3. Combining non-normal copula with normal marginals in (17) provides non-normal CDF.

The copula of the random vector (X_1, X_2) in our example is not normal. That is why this random vector is not normal. Its copula possesses the PDF of the form

$$c(u_1, u_2) = \begin{cases} 2, & (u_1, u_2) \in [0, 1/2] \times [0, 1/2] \text{ or } (u_1, u_2) \in [1/2, 1] \times [1/2, 1], \\ 0, & \text{elsewhere in } [0, 1] \times [0, 1]. \end{cases} \tag{18}$$

In the figure 5.a the region with $c(u_1, u_2) = 2$ is shown in gray.

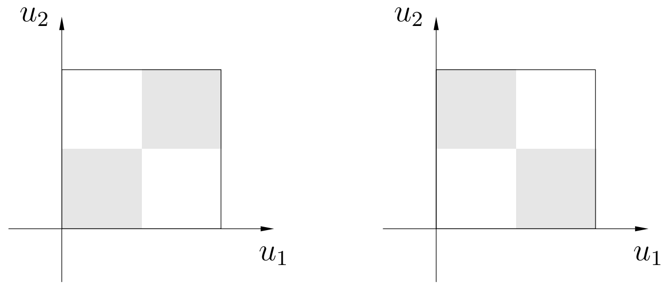
Slightly changing the definition of X_2 from (4) to

$$X_2 = \begin{cases} -|Z|, & X_1 \geq 0, \\ |Z|, & X_1 < 0, \end{cases} \tag{19}$$

we obtain the modification of the example. Its copula $\widetilde{c}$ has the form

$$\widetilde{c}(u_1, u_2) = \begin{cases} 2, & (u_1, u_2) \in [0, 1/2] \times [1/2, 1] \text{ or } (u_1, u_2) \in [1/2, 1] \times [0, 1/2], \\ 0, & \text{elsewhere in } [0, 1] \times [0, 1], \end{cases} \tag{20}$$

and is shown in the figure 5.b. Note that the copula with the PDF c exhibits positive dependence, while the copula with the PDF $\widetilde{c}$ exhibits negative dependence.



a) copula c b) copula $\widetilde{c}$

Figure 5: Areas of positive values of the PDF of the copulas c, $\widetilde{c}$ in (18), (20)

6 Miscellaneous

Let's show using (15) that $F(z) \to 1$ as $z \to \infty$. Since $1 - \Phi(z) \to 0$ as $z \to \infty$, it suffices to show

$$\lim_{z \to \infty} \int_0^z \varphi(x)\Phi(z-x)\,dx = \frac{1}{2}\,. \tag{21}$$

Consider any increasing function $f(z)$, $z \geq 0$ with $f(z) \to \infty$ and $f(z)/z \to 0$ as $z \to \infty$ (say $f(z) = z^{1/2}$). For z sufficiently large we have $f(x) < z$ and

$$\begin{aligned} &\int_0^z \varphi(x)\Phi(z-x)\,dx - \int_0^{f(z)} \varphi(x)\Phi(z-x)\,dx \\ &\quad = \int_{f(z)}^z \varphi(x)\Phi(z-x)\,dx \leq \int_{f(z)}^z \varphi(x)\,dx \\ &\quad \leq (1 - \Phi(f(x)) \to 0 \ \text{ as } \ z \to \infty. \end{aligned} \tag{22}$$

On the other hand, $x \leq f(z)$ implies $z - x \geq z - f(z) \to \infty$ as $z \to \infty$, so $\Phi(z - f(z)) \to 1$ as $z \to \infty$, and

$$\int_0^{f(z)} \varphi(x)\Phi(z-x)\,dx \geq \Phi(z - f(z)) \int_0^{f(z)} \varphi(x)\,dx \longrightarrow \int_0^{\infty} \varphi(x)\,dx = \frac{1}{2}\,. \tag{23}$$

Combining (22) with (23) provides the desired (21).

References

[1] A.N. Shiryaev (1995) *Probability.* Springer, 644 p.

[2] G.A. Holton (2003) *Value-at-Risk: Theory and Practice.* San Diego CA: Academic Press, 405 p.

[3] G.A. Holton (2003) Joint normal distribution.
http://www.riskglossary.com/link/joint_normal_distribution.htm

[4] M. Taranto (2006) Linear combination of normal variables is always distributed normally. *Private communication.*

[5] R. Nelsen (1999) *An introduction to copulas*. Volume 139 of *Lecture Notes in Statistics*. Springer Verlag. Berlin, Heidelberg, New York.

[6] Credit Lyonnais (2001) Copulas.
http://gro.creditlyonnais.fr/content/rd/home_copulas.htm