Copulas with Prescribed Correlation Matrix

Luc Devroye and Gérard Letac

Abstract Consider the convex set R_n of semi positive definite matrices of order n with diagonal $(1, \ldots, 1)$. If μ is a distribution in $\mathbb{R}^n$ with second moments, denote by $R(\mu) \in R_n$ its correlation matrix. Denote by C_n the set of distributions in $[0, 1]^n$ with all margins uniform on $[0, 1]$ (called copulas). The paper proves that $\mu \mapsto R(\mu)$ is a surjection from C_n on R_n if $n \leq 9$. It also studies the Gaussian copulas μ such that $R(\mu) = R$ for a given $R \in R_n$.

1 Foreword

Marc Yor was also an explorer in the jungle of probability distributions, either in discovering a new species, or in landing on an unexpected simple law after a difficult trip on stochastic calculus: we remember his enthousiam after proving that $\left(\int_0^\infty \exp(2B(t) - 2st)dt\right)^{-1}$ is gamma distributed with shape parameter s ('The first natural occurrence of a gamma distribution which is not a chi square!'). Although the authors have been rather inclined to deal with discrete time, common discussions with Marc were about laws in any dimension. Here are some remarks—actually initially coming from financial mathematics—where the beta-gamma algebra (a term coined by Marc) has a role.

2 Introduction

The set of symmetric positive semi-definite matrices $(r_{ij})_{1\leq i,j\leq n}$ of order n such that the diagonal elements r_{ii} are equal to 1 for all $i = 1, \ldots, n$ is denoted by $\mathcal{R}_n$. Given a random variable $(X_1, \ldots, X_n)$ on $\mathbb{R}^n$ with distribution μ such that the second

L. Devroye (✉)
School of Computer Science, McGill University, Montreal, QC, Canada H3A 0G4
e-mail: lucdevroye@gmail.com

G. Letac
Equipe de Statistique et Probabilités, Université de Toulouse, Toulouse, France

C. Donati-Martin et al. (eds.), *In Memoriam Marc Yor - Séminaire de Probabilités XLVII*, Lecture Notes in Mathematics 2137,
DOI 10.1007/978-3-319-18585-9_25

moments of the X_i's exist, its correlation matrix

$$R(\mu) = (r_{ij})_{1\leq i,j\leq n} \in \mathcal{R}_n$$

is defined by r_{ij} as the correlation of X_i and X_j if $i < j$, and $r_{ii} = 1$. A copula is a probability μ on $[0,1]^n$ such that X_i is uniform on $[0,1]$ for $i = 1,\dots,n$ when $(X_1,\dots,X_n) \sim \mu$. We consider the following problem: given $R \in \mathcal{R}_n$, does there exist a copula μ such that $R(\mu) = R$? The aim of this note is to show that the answer is yes if $n \leq 9$. The present authors believe that this limit $n = 9$ is a real obstruction and that for $n \geq 10$ there exists $R \in \mathcal{R}_n$ such that there is no copula μ such that $R(\mu) = R$.

Section 3 gives some general facts about the convex set $\mathcal{R}_n$. Section 4 proves that if $k \geq 1/2$, if $2 \leq n \leq 5$ and if $R \in \mathcal{R}_n$ there exists a distribution μ on $[0,1]^n$ such that

$$X_i \sim \beta_{k,k}(dx) = \frac{1}{B(k,k)} x^{k-1}(1-x)^{k-1}\mathbf{1}_{(0,1)}(x)dx \tag{1}$$

if $(X_1,\dots,X_n) \sim \mu$. This is an extension of the previous statement since $\beta_{k,k}$ is the uniform distribution if $k = 1$. Section 5 proves the remainder of the theorem, namely for $6 \leq n \leq 9$. Section 6 considers the useful and classical Gaussian copulas and explains why there are $R \in \mathcal{R}_n$ that cannot be the correlation matrix of any Gaussian copula. The present paper is both a simplification and an extension of the arXiv paper [1].

3 Extreme Points of $\mathcal{R}_n$

The set $\mathcal{R}_n$ is a convex part of the linear space of symmetric matrices of order n. It is clearly closed and if $R = (r_{ij})_{1\leq i,j\leq n} \in Rc_n$ we have $|r_{ij}| \leq 1$: this shows that $\mathcal{R}_n$ is compact. More specifically, $\mathcal{R}_n$ is in the affine subspace of dimension $n(n-1)/2$ of the symmetric matrices of order n with diagonal $(1,\dots,1)$. Its extreme points have been described in [8]. In particular we have

Theorem 1 *If an extreme point of $\mathcal{R}_n$ has rank r then $r(r+1)/2 \leq n$.*

We vizualize this statement:

r	1	2	3	4	5	...
$\frac{r(r+1)}{2}$	1	3	6	10	15	...

- Case $n = 2$. As a consequence the extreme points of $\mathcal{R}_2$ are of rank one. They are nothing but the two matrices

$$\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}, \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix}.$$

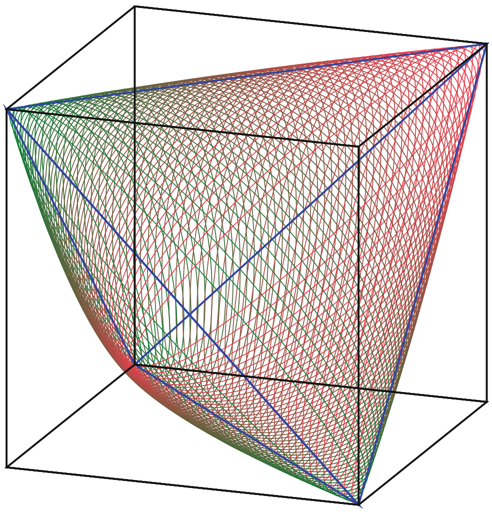

Fig. 1 The space of the three off-diagonal correlation coefficients of a correlation matrix is a convex subset of $[0, 1]^3$

This comes from the fact $R \in \mathcal{R}_2$ of rank one has the form $R = AA^t$ where $A^t = (a_1, a_2)$: since $r_{ii} = 1$ this implies that $a_1^2 = a_2^2 = 1$.

- Case $n \geq 3$. Figure 1 below displays the acceptable values of (x, y, z) when

$$R(x, y, z) = \begin{bmatrix} 1 & z & y \\ z & 1 & x \\ y & x & 1 \end{bmatrix} \tag{2}$$

is positive definite. Its boundary is the part in $|x|, |y|, |z| \leq 1$ of the Steiner surface $1 - x^2 - y^2 - z^2 + 2xyz = 0$.

Proposition 1 *Let $n \geq 3$. Then $R = (r_{ij})_{1 \leq i,j \leq n} \in \mathcal{R}_n$ has rank 2 if and only if there exists n distinct numbers $\alpha_1, \ldots, \alpha_n$ such that $r_{ij} = \cos(\alpha_i - \alpha_j)$.*

Proof $\Rightarrow$: Since R has rank 2 there are two independent vectors A and B of $\mathbb{R}^n$ such that $R = AA^t + BB^t$. Writing $A^t = (a_1, \ldots, a_n)$ and $B^t = (b_1, \ldots, b_n)$ the fact that $r_{ii} = 1$ implies that $a_i^2 + b_i^2 = 1$. Taking $a_i = \cos \alpha_i$ and $b_i = \sin \alpha_i$ gives $r_{ij} = \cos(\alpha_i - \alpha_j)$. $\Leftarrow$: Since only differences $\alpha_i - \alpha_j$ appear in $r_{ij} = \cos(\alpha_i - \alpha_j)$ without loss of generality we take $\alpha_n = 0$ we define $A^t = (\cos \alpha_1, \ldots, \cos \alpha_{n-1}, 1)$ and $B^t = (\sin \alpha_1, \ldots, \sin \alpha_{n-1}, 1)$ and $R = AA^t + BB^t$ is easily checked. $\square$

- Case $n \geq 6$.

Proposition 2 *Let $n \geq 6$. Then $R = (r_{ij})_{1 \leq i,j \leq n} \in \mathcal{R}_n$ has rank 3 if and only if there exist $v_1, \ldots, v_n$ on the unit sphere S_2 of $\mathbb{R}^3$ such that for all $i < j$ we have $r_{ij} = \langle v_i, v_j \rangle$ and such that the system $v_1, \ldots, v_n$ generates $\mathbb{R}^3$.*

Proof The direct proof is quite analogous to Proposition 1: there exist $A, B, C \in \mathbb{R}^n$ such that $R = AA^t + BB^t + CC^t$. and such that A, B, C are independent. Writing

$$[A, B, C] = \begin{bmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ \dots & \dots & \dots \\ a_n & b_n & c_n \end{bmatrix} \tag{3}$$

the desired vectors are $v_i^t = (a_i, b_i, c_i)$. The converse is similar. $\square$

The following proposition explains the importance of the extreme points of $\mathcal{R}_n$ for our problem.

Proposition 3 *Let $X = (X_1, \dots, X_n) \sim \mu$ and $Y = (Y_1, \dots, Y_n) \sim \nu$ be two random variables of $\mathbb{R}^n$ such that for all $i = 1, \dots, n$ we have $X_i \sim Y_i$ and X_i has second moments and are not Dirac. Then for all $\lambda \in (0, 1)$ we have*

$$R(\lambda\mu + (1 - \lambda)\nu) = \lambda R(\mu) + (1 - \lambda)R(\nu).$$

Proof $X_i \sim Y_i$ implies that the mean m_i and the dispersion σ_i of X_i and Y_i are the same. Denote $D = \mathrm{diag}(\sigma_1, \dots, \sigma_n)$. Since the X_i are not Dirac, D is invertible. Denote by

$$\Sigma(\mu) = (\mathbb{E}((X_i - m_i)(X_j - m_j))_{1 \le i,j \le n} = DR(\mu)D$$

the covariance matrix of μ. Define $Z = (Z_1, \dots, Z_n)$ by $Z = X$ with probability λ and $Z = Y$ with probability $(1 - \lambda)$. Thus $Z \sim \lambda\mu + (1 - \lambda)\nu$. Here again the mean and the dispersion of Z_i are m_i and σ_i. Finally the covariance matrix of Z is $\Sigma(\lambda\mu + (1 - \lambda)\nu) = \lambda\Sigma(\mu) + (1 - \lambda)\Sigma(\nu)$ which gives

$$\begin{aligned} R(\lambda\mu + (1 - \lambda)\nu) &= D^{-1}\Sigma(\lambda\mu + (1 - \lambda)\nu)D^{-1} \\ &= \lambda D^{-1}\Sigma(\mu)D^{-1} + (1 - \lambda)D^{-1}\Sigma(\nu)D^{-1} \\ &= \lambda R(\mu) + (1 - \lambda)R(\nu). \end{aligned}$$

$\square$

Corollary 1 *Let $\nu_1, \dots, \nu_n$ a sequence of probabilities on $\mathbb{R}$ having second moments and denote by M the set of probabilities μ on $\mathbb{R}^n$ such that for all $i = 1, \dots, n$ we have $X_i \sim \nu_i$, with $(X_1, \dots, X_n) \sim \mu$. Then the map from M to $\mathcal{R}_n$ defined by $\mu \mapsto R(\mu)$ is surjective if and only if for any extreme point R of $\mathcal{R}_n$ there exists a $\mu \in M$ such that $R = R(\mu)$.*

Proof $\Rightarrow$ comes from the definition. $\Leftarrow$: Since the convex set $\mathcal{R}_n$ has dimension $N = n(n - 1)/2$, the Caratheodory theorem implies that if $R \in \mathcal{R}_n$ then there exists

$N+1$ extreme points $R_0, \dots, R_N$ of $\mathcal{R}_n$ and non negative numbers $(\lambda_i)_{i=0}^N$ of sum 1 such that

$$R = \lambda_0 R_0 + \cdots + \lambda_N R_N.$$

From the hypothesis, for $j = 0, \dots, N$ there exists $\mu_j \in M$ such that $R(\mu_j) = R_j$. Define finally

$$\mu = \lambda_0 \mu_0 + \cdots + \lambda_N \mu_N$$

and apply Proposition 3, we get that $R = R(\mu)$ as desired. □

Comments: With the notation of Corollary 1 and the result of Proposition 3, the map $\mu \mapsto R(\mu)$ from M to $\mathcal{R}_n$ is affine. Consider now the case where for all $i = 1, \dots, n$, the probability ν_i is concentrated on a finite number of atoms. In this particular case M is a polytope, and therefore its image $R(M)$ is a polytope contained in $\mathcal{R}_n$. For $n = 3$ clearly $\mathcal{R}_3$ is not a polytope (see Fig. 1) and therefore there exists a $R \in \mathcal{R}_3$ which is not in $R(M)$: with discrete margins, you cannot reach an arbitrary correlation matrix.

4 The Case $3 \leq n \leq 5$ and the Gasper Distribution

In this section we prove (Proposition 5) that if $\nu_1 = \ldots = \nu_n = \beta_{kk}$ as defined by (1) and with $k \geq 1/2$, if M is defined as in Corollary 1 and if $R \in \mathcal{R}_n$ has rank 2 one can find $\mu \in M$ such that $R = R(\mu)$. The corollary of this Proposition 1 will be that for any $R \in \mathcal{R}_n$ with $3 \leq n \leq 5$ one can find μ such that $R(\mu) = R$ and such that the margins of μ are β_{kk}. Proposition 4 relies on the existence of a special distribution $\Phi_{k,r}$ called the Gasper distribution in the plane that we are going to describe.

Definition Let $k \geq 1/2$. Let $D > 0$ such that $D^2 \sim \beta_{1,k-\frac{1}{2}}$ (if $k > \frac{1}{2}$) and $D \sim \delta_1$ if $k = \frac{1}{2}$. We assume that D is independent of Θ, uniformly distributed on $(0, 2\pi)$. Let $r \in (-1, 1)$ and $\alpha \in (0, \pi)$ such that $r = \cos\alpha$. The Gasper distribution $\Phi_{k,r}$ is the distribution of $(D\cos\Theta, D\cos(\Theta - \alpha))$.

Proposition 4 *If $(X_1, X_2) \sim \Phi_{k,r}$ then X_1 and X_2 have distribution $\nu_k(dx) = \frac{1}{B(k,k)}(1-x^2)^{k-1}1_{(-1,1)}(x)dx$ and correlation r.*

Proof Clearly $X_1 \sim -X_1$ and for seeing that $X_1 \sim \nu_k$ enough is to prove that

$$\mathbb{E}(X_1^{2s}) = \frac{2^{1-2k}}{B(k,k)} \int_{-1}^{1} x^{2s}(1-x^2)^{k-1}dx \tag{4}$$

The right-hand side of (4) is

$$\frac{2^{2-2k}}{B(k,k)}\int_0^1 x^{2s}(1-x^2)^{k-1}dx = 2^{1-2k}\frac{\Gamma(s+\frac{1}{2})\Gamma(2k)}{\Gamma(s+\frac{1}{2}+k)\Gamma(k)}.$$

The left-hand side of (4) is

$$\mathbb{E}(D^{2s})\mathbb{E}((\cos^2\Theta)^s) = \frac{\Gamma(s+1)\Gamma(k+\frac{1}{2})}{\Gamma(s+k+\frac{1}{2})} \times \frac{\Gamma(s+\frac{1}{2})}{\sqrt{\pi}\Gamma(s+1)}.$$

Using the duplication formula $\Gamma(k)\Gamma(k+\frac{1}{2}) = 2^{1-2k}\sqrt{\pi}\,\Gamma(2k)$ proves (4). Since Θ is uniform one has $\cos(\Theta-\alpha) \sim \cos\Theta$ and $X_1 \sim X_2$. For showing that the correlation of (X_1, X_2) is $r = \cos\alpha$ we observe that

$$\mathbb{E}(X_1^2) = \mathbb{E}(D^2)\mathbb{E}(\cos^2\Theta) = \frac{1}{2k+1}$$

$$\mathbb{E}(X_1X_2) = \mathbb{E}(D^2)\mathbb{E}(\cos\Theta\cos(\Theta-\alpha)) = \frac{\cos\alpha}{2k+1}. \square$$

Comments: It is worthwhile to say a few things about this Gasper distribution. It is essentially considered in two celebrated papers by George Gasper [3] and [4]. If $k = \frac{1}{2}$ then $\Phi_{\frac{1}{2},r}$ is concentrated on the ellipse $E_r = E_{cos\alpha}$ parameterized by the circle as

$$\theta \mapsto (x(\theta), y(\theta)) = (\cos\theta, \cos(\theta-\alpha))$$

$$E_r = \{(x,y); (y-xr)^2 = (1-x^2)(1-r^2)\} = \{(x,y); \Delta(x,y,z) = 0\}$$

where

$$\Delta(x,y,r) = \det\begin{bmatrix} 1 & r & y \\ r & 1 & x \\ y & x & 1 \end{bmatrix} = 1 - x^2 - y^2 - r^2 + 2xyr$$

(Compare with (2)). Now denote by $U_r = \{(x,y); \Delta(x,y,r) > 0\}$ the interior of the convex hull of E_r and assume that $k > \frac{1}{2}$. Then Gasper shows that

$$\Phi_{r,k}(dx,dy) = \frac{2k-1}{2\pi}(1-r^2)^{\frac{1}{4}-\frac{k}{2}}\Delta(x,y,r)^{k-\frac{3}{2}}\mathbf{1}_{U_r}(x,y)\,dxdy.$$

The Gasper distribution $\phi_{k,r}$ appears as a Lancaster distribution (see [7]) for the pair (ν_k, ν_k). More specifically consider the sequence $(Q_n)_{n=0}^{\infty}$ of the orthonormal polynomials for the weight ν_k. Thus Q_n is the Jacobi polynomial $P_n^{k-1,k-1}$ normalized such that

$$\int_{-1}^{1} Q_n^2(x)\nu_k(dx) = 1.$$

For $1/2 < k$ denote

$$K(x,y,z) = \sum_{n=0}^{\infty} \frac{Q_n(x)Q_n(y)Q_n(z)}{Q_n(1)}.$$

This series converges if $|x|, |y|, |z| < 1$ and its sum is zero when (x,y) is not in the interior U_r of the ellipse E_r. With this notation we have

$$\phi_{k,r}(dx,dy) = K(x,y,r)\nu_k(dx)\nu_k(dy).$$

This result is essentially due to [3] (with credits to Sonine, Gegenbauer and Moller). See [5, 6] for details.

Proposition 5 *Let $\alpha_1,\dots,\alpha_n$ which are distinct modulo π. Let*

$$R = (\cos(\alpha_i - \alpha_j))_{1\le i,j\le n} \in \mathcal{R}_n$$

and consider the two-dimensional plane $H \subset \mathbb{R}^n$ generated by $c = (\cos\alpha_1,\dots,\cos\alpha_n)$ and $s = (\sin\alpha_1,\dots,\sin\alpha_n)$. Consider the random variable $X = (X_1,\dots,X_n)$ concentrated on H such that $(X_1,X_2) \sim \Phi_{k,\cos(\alpha_1-\alpha_2)}$ and denote by μ the distribution of X. Then

- *For $1 \le i < j \le n$ we have $(X_i,X_j) \sim \Phi_{k,\cos(\alpha_i-\alpha_j)}$*
- *$R = R(\mu)$.*

Proof Recall that $R \in \mathcal{R}_n$ from Proposition 1. Since $X \in H$ there exists A, B such that for all $i = 1,\dots,n$ one has $X_i = A\cos\alpha_i + B\sin\alpha_i$. From the fact that $(X_1,X_2) \sim \Phi_{k,\cos(\alpha_1-\alpha_2)}$ we can claim the existence of a (Θ, D) such that Θ is uniform on the circle and is independent of $D > 0$ such that $D^2 \sim \beta_{1,k-\frac{1}{2}}$ and such that

$$(X_1,X_2) \sim D\cos(\Theta-\alpha_1), D\cos(\Theta-\alpha_2)).$$

From an elementary calculation this leads to saying that $(A,B) \sim (D\cos\Theta, D\sin\Theta)$ and finally that

$$(X_1,\dots,X_n) \sim (D\cos(\Theta-\alpha_1),\dots,D\cos(\Theta-\alpha_n)).$$

From Proposition 4 this proves the results. □

Conclusion: The previous proposition has shown that for $k \geq \frac{1}{2}$ and for any extremal point R of $\mathcal{R}_n$ there exists a distribution μ_R in $(-1,1)^n$ with margins ν_k and correlation matrix R. From Corollary 1 above, since an arbitrary $R \in \mathcal{R}_n$ is a convex combination $R = \lambda_0 R_0 + \cdots + \lambda_n R_n$ of extreme points R_i of $\mathcal{R}_n$ the distribution $\mu = \lambda_0 \mu_{R_0} + \cdots + \lambda_n \mu_{R_n}$ has margins ν_k and correlation R.

Since ν_k is the affine transformation of $\beta_{k,k}$ by $u \mapsto x = 2u - 1$ this implies that there exists also a distribution in $(0,1)^k$ with margins $\beta_{k,k}$ and correlation matrix R. Since $\beta_{1,1}$ is the uniform distribution on $(0,1)$ a corollary is the existence of a copula with arbitrary correlation matrix R.

Example To illustrate Proposition 5 consider the case $n = 3$ and $R \in \mathcal{R}_3$ defined by

$$R = \begin{bmatrix} 1 & -\frac{1}{2} & -\frac{1}{2} \\ -\frac{1}{2} & 1 & -\frac{1}{2} \\ -\frac{1}{2} & -\frac{1}{2} & 1 \end{bmatrix}$$

which is an extreme point corresponding to $\alpha_1 = 0, \alpha_2 = 2\pi/3 = -\alpha_3$. This example is important since, as we are going to observe in Sect. 6, it is not possible to find a Gaussian copula having R as correlation matrix. Recall now a celebrated result:

Archimedes Theorem: If X is uniformly distributed on the unit sphere S of the three-dimensional Euclidean space E and if Π is an orthogonal projection of E on a one-dimensional line $F \subset E$ then $\Pi(X)$ is uniform on the diameter with end points $S \cap F$.

Proof While we learnt a different proof in 'classe de Première' in the middle of the fifties, here is a computational proof: let $Z \sim N(0, \mathrm{id}_E)$. Then $X \sim Z/\|Z\|$. Choose orthonormal coordinates (x_1, x_2, x_3) such that F is the x_1 axis. As a consequence of $Z = (Z_1, Z_2, Z_3)$ we have $X_1^2 \sim Z_1^2/(Z_1^1 + Z_2^2 + Z_3^2)$ and since the Z_i^2 are chi square independent with one degree of freedom, this implies that $X_1^2 \sim \beta_{1/2,1}$ which leads quickly to X_1 uniformly distributed on $(-1,1)$since $X_1 \sim -X_1$. $\square$

Proposition 5 offers a construction (see Fig. 2) of a distribution in $C = [-1,1]^3$ with uniform margins ν_1 on $(-1,1)$ as a distribution concentrated on the plane P of equation $x + y + z = 1$. The intersection $C \cap P$ is a regular hexagon. Introduce the disc D inscribed in the hexagon $C \cap P$ and the sphere S admitting the boundary of D as one of its grand circles. Now consider the uniform distribution on S. Denote by μ its orthogonal projection μ on D. Actually any orthogonal projection of μ on a diameter of D is uniform on this diameter, from Archimedes Theorem. Apply this to the three diagonals of the hexagon $C \cap P$: this proves that the three margins of μ are the uniform measure ν_1.

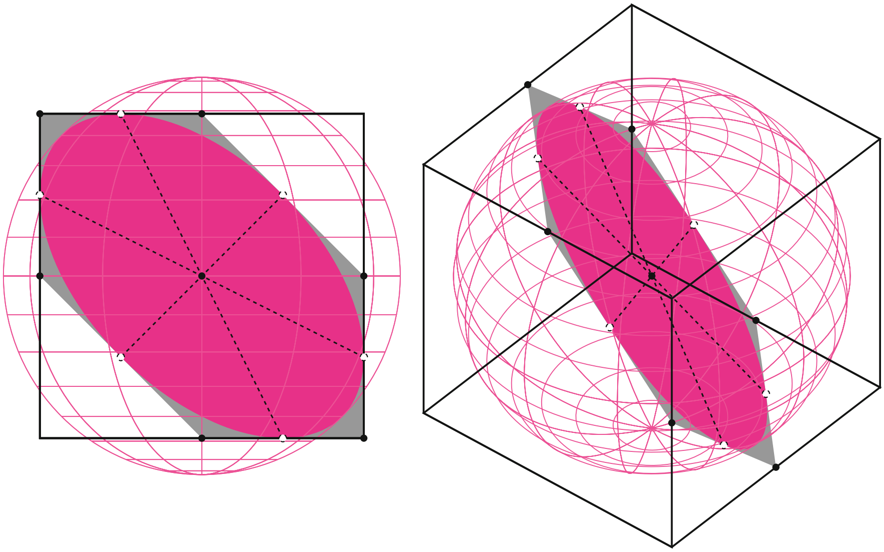

Fig. 2 Illustration of our construction. First take a point uniformly on the surface of the ball. Project it to the plane shown (so that it falls in the *circle*). The three coordinates of that point are each uniformly distributed on $[-1, 1]$

5 The Case $6 \leq n \leq 9$

Proposition 6 *Let $n \geq 6$ and let A, B, C be three independent vectors of $\mathbb{R}^n$ such that $R = [A, B, C][A^t, B^t, C^t]^t = AA^t + BB^t + CC^t$ is a correlation matrix. Let $Y = (U, V, W)$ be uniformly distributed on the unit sphere $S_2 \subset \mathbb{R}^3$ and let μ be the distribution of $X = AU + BV + CW$ in $\mathbb{R}^n$. Then $R(\mu) = R$ and the marginal distributions of μ are ν_1, the uniform distribution in* $(-1, 1)$.

Proof From Archimedes Theorem, U, V and W have distribution ν_1. Furthermore, since the distribution of (U, V) is invariant by rotation, then $(U, V) \sim (D\cos\Theta, D\sin\Theta)$ where $D = \sqrt{U^2 + V^2}$ is independent of Θ uniform on the circle. This implies that $\mathbb{E}(UV) = 0$. Since $\mathbb{E}(U^2) = 1/3$ the covariance matrix of (U, V, W) is $I_3/3$. From this remark, and using the fact that $AU + BV + CW$ is centered, the covariance matrix of $AU + BV + CW$ is

$$\mathbb{E}((AU + BV + CW)(AU + BV + CW)^t) = R/3$$

and this proves $R(\mu) = R$. Finally, using the representation (4) of the matrix $[A, B, C]$ and denoting $v_i = (a_i, b_i, c_i)$ we see that the component X_i of $AU + BV + CW$ is $a_iU + b_iV + c_iW = \langle v_i, Y\rangle$. Since $\|v_i\|^2 = 1$ the random variable X_i is the orthogonal projection of Y on $\mathbb{R}v_i$ and is uniform on $(-1, 1)$ from Archimedes Theorem. □

Comments: The above proposition finishes the proof of the fact that for $n \leq 9$, and if R is an extreme point of $\mathcal{R}_n$ then it is the correlation of some copula. From Proposition 3 this completes the proof that any $R \in \mathcal{R}_n$ is the correlation of a copula for $n \leq 9$. The fact that this result can be extended to $n \geq 10$ is doubtful, since there are $R \in \mathcal{R}_{10}$ of the form $AA^t + BB^t + CC^t + DD^t$ where $A, B, C, D \in \mathbb{R}^{10}$ and the technique of the proof of Proposition 6 seems to indicate that it is impossible. A similar phenomenon seems to occur if we want to construct a distribution μ in $\mathbb{R}^6$ such that $R(\mu)$ has rank 3 and such that the margins of μ are $\beta_{1/2,1/2}$.

Accordingly, we conjecture the existence of $R \in \mathcal{R}_{10}$ which cannot be the correlation of a copula, and we conjecture the existence of $R \in \mathcal{R}_6$ which cannot be the correlation of a distribution whose margins are the arsine distribution.

6 Gaussian Copulas

In this section, we explore the simplest idea for building a copula on $[0, 1]^n$ with a non trivial variance: select a Gaussian random variable $(X_1, \dots, X_n) \sim N(0, R)$ where $R \in \mathcal{R}_n$, introduce the distribution function

$$\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-t^2/2} dt$$

of $N(0, 1)$ and observe that the law μ of $(U_1, \dots, U_n) = (\Phi(X_1), \dots, \Phi(X_n))$ is a copula. A μ which can be obtained in that way is called a Gaussian copula. However its correlation $R^* = R(\mu)$ is not equal to R except in trivial cases.

Therefore this section considers the map from $\mathcal{R}_n$ to itself defined by $R \mapsto R^*$. This map is not surjective: in particular, in comments following Proposition 7 we exhibit a correlation matrix which cannot be the correlation of a Gaussian copula. First we compute R^* by brute force (Proposition 7), getting a result of [2]. We make also two remarks about the expectation of $f_1(X)f_2(Y)$ when (X, Y) is centered Gaussian (Propositions 8 and 9). Proposition 10 leads to a more elegant proof of Proposition 7 by using Hermite polynomials.

Proposition 7 *Let $R = (r_{ij})_{1 \leq i,j \leq n}$ be a correlation matrix, let*

$$(X_1, \dots, X_n) \sim N(0, R)$$

and let μ be the law of $(U_1, \dots, U_n) = (\Phi(X_1), \dots, \Phi(X_n))$. Then

$$R(\mu) = R^* = (g(r_{ij}))_{1 \leq i,j \leq n}$$

where

$$g(r) = \frac{6}{\pi} \arcsin \frac{r}{2}. \tag{5}$$

Proof We begin with a standard calculation. We start with (X, Y) centered Gaussian with covariance

$$\Sigma_r = \begin{bmatrix} 1 & r \\ r & 1 \end{bmatrix}. \tag{6}$$

We now compute the quadruple integral

$$f(r) = \mathbb{E}(\Phi(X)\Phi(Y)) = \int_{\mathbb{R}^4} e^{-\frac{1}{2}(u^2+v^2+\frac{1}{1-r^2}(x^2-2rxy+y^2))} \mathbf{1}_{u<x,v<y} \frac{dxdydudv}{(2\pi)^2\sqrt{1-r^2}}.$$

Performing the change of variables $(x, y, u, v) \mapsto (x, y, x-u, y-v) = (x, y, t, s)$ we get

$$f(r) = \frac{1}{\sqrt{4-r^2}} \int_0^\infty \int_0^\infty e^{-\frac{1}{2}(t^2+s^2)} g(r, t, s) \frac{dtds}{2\pi}$$

with

$$g(r, t, s) = \sqrt{\frac{4-r^2}{1-r^2}} \int_{\mathbb{R}^2} e^{xt+ys-\frac{1}{2(1-r^2)}((2-r^2)x^2-2rxy+(2-r^2)y^2)} \frac{dxdy}{2\pi}.$$

Consider

$$A = \frac{1}{1-r^2} \begin{bmatrix} 2-r^2 & -r \\ -r & 2-r^2 \end{bmatrix}, \; B = \frac{1}{4-r^2} \begin{bmatrix} 2-r^2 & r \\ r & 2-r^2 \end{bmatrix}.$$

Then $B = A^{-1}$, $\det A = \frac{4-r^2}{1-r^2}$ and $\det B = \frac{1-r^2}{4-r^2}$. Therefore $g(r, t, s)$ is the Laplace transform of a centered random Gaussian random variable with covariance matrix B. We get

$$g(r, t, s) = e^{\frac{1}{2(4-r^2)}((2-r^2)t^2+2rts+(2-r^2)s^2)}$$

and therefore

$$f(r) = \frac{1}{\sqrt{4-r^2}} \int_0^\infty \int_0^\infty e^{-\frac{1}{2(4-r^2)}(2t^2-2rts+2s^2)} \frac{dtds}{2\pi}.$$

Now we use the fact that if (T, S) is a Gaussian centered random variable with correlation coefficient $\cos\alpha$ with $0 < \alpha < \pi$ then $\Pr(T > 0, S > 0)$ is explicit. For computing it, just introduce $S' \sim N(0,1)$ independent of T observe that $(T,S) \sim (T, T\cos\alpha + S'\sin\alpha)$ and finally write $(T,S') = (D\cos\Theta, D\sin\Theta)$ where $D > 0$ and Θ are independent and where Θ is uniform on $(0,2\pi)$. This leads to

$$\Pr(T > 0, S > 0) = \Pr(\cos\Theta > 0,\ \cos(\Theta-\alpha) > 0) = \frac{\pi-\alpha}{2\pi}.$$

We apply this principle to the above integral which can be seen as

$$\Pr(T > 0, S > 0) = f(r)$$

when $(T,S) \sim N(0, \begin{bmatrix} 2 & r \\ r & 2 \end{bmatrix})$. The correlation coefficient of (T,S) is here $\cos\alpha = \frac{r}{2}$ and we finally get

$$f(r) = \frac{1}{2\pi}(\pi - \arg\cos\frac{r}{2}) = \frac{1}{2} - \frac{1}{2\pi}\arg\cos\frac{r}{2}.$$

Now we consider the function

$$g(r) = 12\mathbb{E}((\Phi(X) - 1/2)(\Phi(Y) - 1/2)) = 12f(r) - 3 = \frac{6}{\pi}\arg\sin\frac{r}{2}$$

and the function $T(x) = 2\sqrt{3}(\Phi(x) - 1/2)$. Thus the random variables $T(X)$ and $T(Y)$ are uniform on $(-\sqrt{3}, \sqrt{3})$ with mean 0, variance 1 and correlation $g(r)$. This implies that the correlation between $\Phi(X)$ and $\Phi(Y)$ is $g(r)$. Coming back to the initial $(X_1,\ldots,X_n)$ the correlation between $\Phi(X_i)$ and $\Phi(X_j)$ is $g(r)$. □

Comments: The function g is odd and increasing since $g'(r) = \frac{6}{\pi\sqrt{4-r^2}}$. Thus we have $|g(r)| < r < 1$. It satisfies $g(0) = 0$, $g(\pm 1) = \pm 1$, $g'(0) = \frac{3}{\pi}$ and $g'(1) = \frac{2\sqrt{3}}{\pi}$. Finally for $-1 < \rho < 1$ we have

$$\rho = g(r) \Leftrightarrow r = 2\sin\frac{\pi\rho}{6}.$$

Calculation shows that for $-1 < \rho < 1$ we have $0 \le |2\sin\frac{\pi\rho}{6} - \rho| \le 0.0180\ldots$ therefore the two functions are quite close. It is useful to picture g and its inverse function in Fig. 3. Observe also that if $\rho = -1/2$ we get

$$r = -2\sin\frac{\pi}{12} = -\frac{\sqrt{3}-1}{\sqrt{2}} = -0.51\ldots < -1/2.$$

An important consequence is the fact that since $r < -1/2$ the matrix $R(r,r,r)$ of (2) is not a correlation matrix and therefore the correlation matrix $R(-\frac{1}{2},-\frac{1}{2},-\frac{1}{2})$ cannot be the correlation matrix of a Gaussian copula. Falk [2] makes essentially a similar observation.

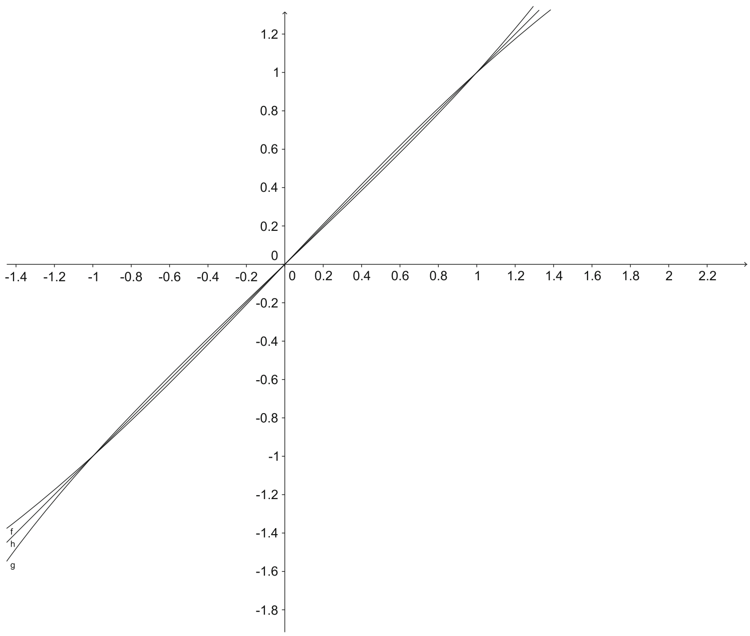

Fig. 3 Graphs of $\rho = g(r) = \frac{6}{\pi} \arg\sin \frac{r}{2}$ and $r = g^{-1}(\rho) = 2 \sin \frac{\pi\rho}{6}$

In the sequel, we proceed to a more general study of the correlation between $f_1(Y_1)$ and $f_2(Y_2)$ when $(Y_1, Y_2) \sim N(0, \Sigma_r)$ as defined in (6). We thank Ivan Nourdin for a shorter proof of the following proposition:

Proposition 8 *Given any $r \in [-1, 1]$ consider the Gaussian random variable $(Y_1, Y_2) \sim N(0, \Sigma_r)$. Consider two probabilities ν_1 and ν_2 on $\mathbb{R}$ with respective distribution functions G_1 and G_2. Then the correlation of $G_1(Y_1)$ and $G_2(Y_2)$ is a continuous increasing function of r.*

Proof We use the fact that if $f \in C^2(\mathbb{R}^2)$ then

$$\frac{d}{dr}\mathbb{E}(f(Y_1, Y_2)) = \mathbb{E}(\frac{\partial^2}{\partial y_1 \partial y_2} f(Y_1, Y_2)) \tag{7}$$

To see this recall that if $X \sim N(0, 1)$ then an integration by parts gives

$$\mathbb{E}(X\varphi(X)) = \mathbb{E}(\varphi'(X)). \tag{8}$$

Writing $Y_2 = rY_1 + \sqrt{1-r^2}Y_3$ where Y_1 and Y_3 are independent $N(0,1)$ we get

$$\frac{d}{dr}\mathbb{E}(f(Y_1,Y_2)) = \mathbb{E}((Y_1 - \frac{r}{\sqrt{1-r^2}}Y_3)\frac{\partial}{\partial y_2}f(Y_1,Y_2)) \tag{9}$$

$$= \mathbb{E}(Y_1\frac{\partial}{\partial y_2}f(Y_1,Y_2)) - \frac{r}{\sqrt{1-r^2}}\mathbb{E}(Y_3\frac{\partial}{\partial y_2}f(Y_1,Y_2))$$

$$= \mathbb{E}(Y_1\frac{\partial}{\partial y_2}f(Y_1,Y_2)) - r\mathbb{E}(\frac{\partial^2}{\partial y_2^2}f(Y_1,Y_2)) \tag{10}$$

$$= \mathbb{E}(\frac{\partial^2}{\partial y_1 \partial y_2}f(Y_1,Y_2)) \tag{11}$$

In this sequence of equalities (9) is derivation inside an integral, (10) is the application of (8) to $\varphi(Y_3) = \frac{\partial}{\partial y_2}f(Y_1, rY_1 + \sqrt{1-r^2}Y_3))$ and (11) is the application of (8) to $\varphi(Y_1) = \frac{\partial}{\partial y_2}f(Y_1, rY_1 + \sqrt{1-r^2}Y_3))$ which satisfies

$$\varphi'(Y_1) = \frac{\partial^2}{\partial y_1 \partial y_2}f(Y_1,Y_2) + r\frac{\partial^2}{\partial y_2^2}f(Y_1,Y_2).$$

The application of (7) to the proof of Proposition 1 is clear: if G_1 and G_2 are smooth enough, we take $f(y_1,y_2)$ as $G_1(y_1)G_2(y_2)$. If not we use an approximation. □.

Corollary 2 *Given two probability distributions μ_1 and μ_2 on the real line having second moments with respective distribution functions F_1 and F_2. Given any $r \in [-1,1]$ consider the Gaussian random variable $(Y_1,Y_2) \sim N(0,\Sigma_r)$. Then $(X_1,X_2) = F_1^{-1}(\Phi(Y_1)), F_1^{-1}(\Phi(Y_2))$ has a correlation*

$$\rho = g_{\mu_1,\mu_2}(r)$$

which is a continuous increasing function on $[-1,1]$. In particular if $g_{\mu_1,\mu_2}(-1) = a$ and $g_{\mu_1,\mu_2}(1) = b$ and if $a \le \rho \le b$ there exists a unique $r = f_{\mu_1,\mu_2}(\rho) \in [-1,1]$ such that (X_1,X_2) has correlation ρ.

Proposition 9 *Let (X,Y) be a centered Gaussian variable of $\mathbb{R}^2$ with covariance matrix Σ_r and let $f : \mathbb{R} \to \mathbb{R}$ be a function such that $\mathbb{E}(f(X)) = 0$ and $\mathbb{E}(f(X)^2) = 1$. Then $\mathbb{E}(f(X)f(Y)) = r$ for all $-1 \le r \le 1$ if and only if $f(x) = \pm x$.*

Proof Write $r = \cos\alpha$ with $0 \le \alpha \le \pi$. If X, Z are independent centered real Gaussian random variables with variance 1, then $Y = X\cos\alpha + Z\sin\alpha$ is centered with variance 1, (X,Y) is Gaussian and $\mathbb{E}(XY) = \cos\alpha$. Therefore we rewrite this as

$$\cos\alpha = \int_{\mathbb{R}^2} f(x)f(x\cos\alpha + z\sin\alpha)e^{-\frac{1}{2}(x^2+z^2)}\frac{dxdz}{2\pi} \tag{12}$$

$$rangle = \int_0^\infty \rho e^{-\frac{\rho^2}{2}}\left(\frac{1}{2\pi}\int_{-\pi}^{\pi} f(\rho\cos\theta)f(\rho\cos(\alpha-\theta))d\theta\right)d\rho \tag{13}$$

where we have used polar coordinates $x = \rho\cos\theta$ and $z = \rho\sin\theta$ for the second equality. This equality is established for $0 \le \alpha \le \pi$ but it is still correct when we change α into $-\alpha$. Now we introduce the Fourier coefficients for n in the set $\mathbb{Z}$ of relative integers:

$$\hat{f}_n(\rho) = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(\rho\cos\theta)e^{-in\theta}d\theta.$$

Since f is real we have the Hermitian symmetry $\hat{f}_{-n}(\rho) = \overline{\hat{f}_n(\rho)}$. Expanding the periodic function (13) in Fourier series and considering the Fourier coefficients of $\alpha \mapsto \cos\alpha$ we get for $n \neq \pm 1$

$$\int_0^\infty \rho e^{-\frac{\rho^2}{2}}\hat{f}_n^2(\rho)d\rho = 0 \tag{14}$$

and $\int_0^\infty \rho e^{-\frac{\rho^2}{2}}\hat{f}_{\pm 1}^2(\rho)d\rho = \frac{1}{2}$. Hermitian symmetry implies that $\hat{f}_0^2(\rho)$ is real and since $\int_0^\infty \rho e^{-\frac{\rho^2}{2}}\hat{f}_0^2(\rho)d\rho = 0$ we get that $\hat{f}_0^2(\rho) = 0$ for almost all $\rho > 0$. This is saying that for almost all $\rho > 0$ we have

$$\int_{-\pi}^{\pi} f(\rho\cos\theta)d\theta = 0. \tag{15}$$

Since $\theta \mapsto f(\rho\cos\theta)$ is a real even function we have

$$f(\rho\cos\theta) \sim \sum_{n=1}^{\infty} a_n(\rho)\cos n\theta$$

and the real number $a_n(\rho)$ is equal to $2\hat{f}_n(\rho)$ and to $2\hat{f}_{-n}(\rho)$ which are therefore real numbers. Using (14) they are zero for all $n \neq \pm 1$ and we get almost everywhere that $f(\rho\cos\theta) = a_1(\rho)\cos\theta$ or $f(\rho u) = a_1(\rho)u$ for all $-1 \le u \le 1$. To conclude we write

$$a_1(\rho)u = f(\rho u) = f(\rho_1\frac{\rho}{\rho_1}u) = a_1(\rho_1)\frac{\rho}{\rho_1}u$$

where u is small enough such that $|\frac{\rho}{\rho_1}u| \le 1$. This implies $\frac{a_1(\rho)}{\rho} = \frac{a_1(\rho_1)}{\rho_1}$ which is a constant c by the principle of separation of variables. Therefore $f(x) = cx$ almost everywhere and $\mathbb{E}(f(X)^2) = 1$ implies that $c = \pm 1$. $\square$

For computing expressions like $\mathbb{E}(f_1(Y_1)f_2(Y_2))$ when $(Y_1, Y_2) \sim N(0, \Sigma_r)$ we use the classical fact below:

Proposition 10 *Let* $(Y_1, Y_2) \sim N(0, \Sigma_r)$. *Let* f_1 *and* f_2 *be real measurable functions such that* $\mathbb{E}(f_i(Y_i)^2)$ *is finite for* $i = 1, 2$. *Consider the Hermite polynomials* $(H_k)_{k=0}^{\infty}$

defined by the generating function

$$e^{xt-\frac{t^2}{2}} = \sum_{k=0}^{\infty} H_k(x)\frac{t^k}{k!}$$

and the expansions

$$f_1(x) = \sum_{k=1}^{\infty} a_k \frac{H_k(x)}{\sqrt{k!}},\ f_2(x) = \sum_{k=1}^{\infty} b_k \frac{H_k(x)}{\sqrt{k!}}.$$

Then for all $-1 \leq r \leq 1$

$$\mathbb{E}(f_1(Y_1)f_2(Y_2)) = \sum_{k=1}^{\infty} a_k b_k r^k.$$

Proof Let us compute

$$\mathbb{E}(e^{Y_1 t-\frac{t^2}{2}} e^{Y_2 s-\frac{s^2}{2}}) = \sum_{k=0}^{\infty}\sum_{m=0}^{\infty} \frac{t^k}{k!}\frac{s^m}{m!}\mathbb{E}(H_k(Y_1)H_m(Y_2)).$$

For this, we use the usual procedure and first write $r = \cos\theta$ with $0 \leq \theta \leq \pi$. If Y_1, Y_3 are independent centered real Gaussian random variables with variance 1, then $Y_2 = Y_1\cos\theta + Y_3\sin\theta$ is centered with variance 1, (Y_1, Y_2) is Gaussian and $\mathbb{E}(Y_1Y_2) = \cos\theta$. Furthermore a simple calculation using the definition of Y_2 gives

$$\mathbb{E}(e^{Y_1 t-\frac{t^2}{2}} e^{Y_2\, s-\frac{s^2}{2}}) = e^{ts\cos\theta}$$

This shows that $\mathbb{E}(H_k(Y_1)H_m(Y_2)) = 0$ if $k \neq m$ and that $\mathbb{E}(H_k(Y_1)H_k(Y_2)) = k!\cos^k\theta$. From this we get the result. □

Corollary 3 *Let* $p_n \geq 0$ *such that* $\sum_{n=1}^{\infty} p_n = 1$ *and consider the generating function* $g(r) = \sum_{n=1}^{\infty} p_n r^n$. *Let* $R = (r_{ij})_{1\leq i,j\leq d}$ *in* $\mathcal{R}_n$. *Then* $R^* = (g(r_{ij}))_{1\leq i,j\leq d}$ *is the covariance matrix of the random variable* $(f(X_1), \ldots, f(X_d))$ *where* $(X_1, \ldots, X_d)$ *is centered Gaussian with covariance R and where*

$$f(x) = \sum_{n=1}^{\infty} \epsilon_n \sqrt{p_n}\frac{H_n(x)}{\sqrt{n!}}$$

with fixed $\epsilon_n = \pm 1$.

Example We have seen an example of such a function f with $f(x) = T(x) = 2\sqrt{3}(\Phi(x) - 1/2)$ and

$$g(r) = \frac{6}{\pi}\arg\sin\frac{r}{2} = \frac{3}{\pi}\sum_{n=0}^{\infty}(\frac{1}{2})_n\frac{1}{4^n n!}\frac{r^{2n+1}}{2n+1}.$$

Thus $p_{2n+1} = \frac{3}{\pi}(\frac{1}{2})_n \frac{1}{4^n n!}\frac{1}{2n+1}$ and $p_{2n} = 0$. For computing ϵ_n we have really to compute

$$\epsilon_n \frac{\sqrt{p_n}}{\sqrt{n!}} = \mathbb{E}(T(X)\frac{H_n(X)}{n!})$$

For this we watch the coefficient of t^n in the power expansion of

$$\mathbb{E}(T(X)e^{Xt-\frac{t^2}{2}})$$

For this we need

$$\mathbb{E}(\Phi(X)e^{Xt-\frac{t^2}{2}}) = 1-\Phi(-\frac{t}{\sqrt{2}}) = \frac{1}{2}+\frac{1}{2\sqrt{\pi}}\sum_{n=0}^{\infty}\frac{(-1)^n}{4^n n!}\frac{t^{2n+1}}{2n+1}$$

$$\mathbb{E}(T(X)e^{Xt-\frac{t^2}{2}}) = \sqrt{\frac{3}{\pi}}\sum_{n=0}^{\infty}\frac{(-1)^n}{4^n n!}\frac{t^{2n+1}}{2n+1}$$

Therefore

$$\epsilon_{2n+1}\frac{\sqrt{p_{2n+1}}}{\sqrt{(2n+1)!}} = \sqrt{\frac{3}{\pi}}\frac{(-1)^n}{4^n n!}\frac{1}{2n+1}$$

which shows that $\epsilon_{2n+1} = (-1)^n$.

References

1. L. Devroye, G. Letac, Copulas in three dimensions with prescribed correlations (2010) [arXiv 1004.3146]
2. M. Falk, A simple approach to the generation of uniformly distributed random variables with prescribed correlations. Commun. Stat. Simul. Comp. **28**, 785–791 (1999)
3. G. Gasper, Positivity and the convolution structure for Jacobi series. Ann. Math. **93**, 112–118 (1971)
4. G. Gasper, Banach algebra for Jacobi series and positivity of a kernel. Ann. Math. **95**, 261–280 (1972)
5. A.E. Koudou, *Problèmes de marges et familles exponentielles naturelles.* Thèse, Université Paul Sabatier, Toulouse, 1995
6. A.E. Koudou, Probabilités de Lancaster. Expositiones Math. **14**, 247–275 (1996)
7. G. Letac, Lancaster probabilities and Gibbs sampling. Stat. Sci. **23**, 187–191 (2008)
8. B. Ycart, Extreme points in convex sets of symmetric matrices. Proc. Am. Math. Soc. **95**(4), 607–612 (1985)