Huber and Marić *Journal of Statistical Distributions and Applications* (2019) 6:2
https://doi.org/10.1186/s40488-019-0091-5

Journal of Statistical Distributions and Applications

RESEARCH Open Access

Admissible Bernoulli correlations

Mark Huber[1†] and Nevena Marić[2*†]

*Correspondence:
maric@math.umsl.edu
†Mark Huber and Nevena Marić contributed equally to this work.
[2]University of Missouri-St. Louis, 1 University Blvd, 311 Express Scripts Hall, St. Louis 63121-4400, MO, USA
Full list of author information is available at the end of the article

Abstract
A multivariate symmetric Bernoulli distribution has marginals that are uniform over the pair $\{0, 1\}$. Consider the problem of sampling from this distribution given a prescribed correlation between each pair of variables. Not all correlation structures can be attained. Here we completely characterize the admissible correlation vectors as those given by convex combinations of simpler distributions. This allows us to bijectively relate the correlations to the well-known CUT_n polytope, as well as determine if the correlation is possible through a linear programming formulation.

Keywords: Bernoulli distribution, Extreme correlations, CUT polytope

Mathematics Subject Classification (2000): 62H20, 60E05, 52B12

Introduction

Consider the admissible correlations among n random variables $(X_1, \ldots, X_n)$ for given marginal distributions. This topic has a long history, dating back to de Finetti (1937) where the problem of maximum negative achievable correlation among n random variables was studied. Fréchet (1951) and Hoeffding (1940) studied the general form of the question, which grew out questions posed by Lévy (1937).

The big question is: can we completely describe set of correlation matrices for a given set of marginal distributions? When $n = 2$ the answer is completely known in terms of Fréchet-Hoeffding bounds. This two dimensional problem was also studied in (Leonov and Qaqish B) for a wide range of distributions.

Therefore we consider dimensions greater than two here. We show for general marginals that if a particular vector calculated from the target correlations and marginals falls into the CUT_n polytope (the convex hull of cut vectors in a complete graph with vertices $\{1, \ldots, n\}$), then there does exist such a joint distribution. This condition is both necessary and sufficient in the case of symmetric Bernoulli marginals.

Correlation matrices are symmetric positive semi-definite and have all ones on the diagonal, denote this set of matrices (of size n by n) as $\mathcal{E}_n$. This convex compact set is called the *elliptope* (see Laurent and Poljak 1995).

For Gaussian marginals, the entirety of $\mathcal{E}_n$ is admissible as correlations, but this is the only nontrivial set of marginals for which the question has been settled. Even for other common distributions surprisingly little is known. One case that has been partially explored is that of copulas. A probability measure on $[0, 1]^n$ is a *copula* if all its marginals are uniformly distributed on $[0, 1]$. Devroye and Letac (2015) have shown that every element in $\mathcal{E}_n$ is a correlation matrix for some copula, for $n \le 9$, but they believe that the statement does not hold for $n \ge 10$.

Here we focus on symmetric Bernoulli variables, that is marginals X_i where $\mathbb{P}(X_i = 1) = \mathbb{P}(X_i = 0) = 1/2$. (Write $X_i \sim \mathsf{Bern}(1/2)$). In Huber and Marić (Huber and Marić 2015) this distribution was shown to be in a certain sense the most difficult marginal: for general marginals it is often possible to transform the problem into symmetric Bernoulli marginals.

This problem, in different guises, appears in numerous fields: physics (Smith and Adelfang 1981), engineering (Lampard 1968), ecology (dos Santos Dias et al. 2008), and finance (Lawrance and Lewis 1981), to name just a few. Due to its applicability in the generation of synthetic optimization problems, it has also received special attention by the simulation community (Hill and Reilly 1994; Henderson et al. 2000).

It should be noted that the answer for symmetric Bernoulli marginals will be a strict subset of $\mathcal{E}_n$, even when n is small. As a simple example consider

$$\begin{pmatrix} 1 & -0.4 & -0.4 \\ -0.4 & 1 & -0.4 \\ -0.4 & -0.4 & 1 \end{pmatrix}.$$

While this matrix is in the elliptope $\mathcal{E}_3$, it cannot be the correlation matrix of three random variables with symmetric Bernoulli marginals. This follows from the results given in the next section (see also Huber and Marić 2015).

Let us note also that knowing the admissible correlations allows us to place the correlation estimates in perspective, which is of great significance in empirical data analysis. Chaganty and Joe (2006) write about errors caused by the belief that any matrix in $\mathcal{E}_n$ is a possible correlation matrix for a set of binary random variables. In the same paper they were able to characterize the achievable correlation matrices when the marginals are Bernoulli. When the dimension is 3 their characterization is easily checkable (as for the 3 by 3 matrix given above), in higher dimensions they give a number of inequalities that grows exponentially in the dimension. They also give an approximate method for checking attainability of the correlation matrix in higher dimensions.

In this paper we give a complete characterization of the correlation matrices for multivariate symmetric Bernoulli distributions by explicitely identifying vertices of the corresponding polytope. This approach leads also to a novel sampling method from the desired marginals and correlations.

The rest of the paper is organized as follows. In the next section it is shown that the question of admissible correlations of multivariate symmetric Bernoulli random variables can be reduced to a subset of distributions that has even more symmetry. This also allows us to bijectively relate the admissible correlations to the well-known CUT_n polytope. In the following section this idea is then used to give a method for construction of a multivariate exponential distribution with prescribed correlation structure. In the last section we discuss our findings in a larger context.

The main result

Consider a vertex of the n-dimensional cube $v \in \{0,1\}^n$. For instance, when $n = 5$, $v = (0,0,1,0,1)$ is such a vertex. Let $\mathbf{1}$ denote the vector of all 1's. Then for any $v \in \{0,1\}^n$, the distribution $\mathsf{Unif}(\{v, \mathbf{1} - v\})$ (discrete uniform distribution over two points: v and $\mathbf{1} - v$) has marginals that are all uniform over the pair $\{0,1\}$. Hence all such distributions are multivariate symmetric Bernoulli.

Any convex combination of multivariate symmetric Bernoulli distributions will also be multivariate symmetric Bernoulli. Our main result is that any admissible correlation structure can also be realized as the correlation structure of such a convex combination.

Theorem 1 *Let ρ be the correlation structure for a multivariate symmetric Bernoulli distribution P. Then there exists P' that is the convex combination of distributions of the form* $\mathsf{Unif}(\{v, \mathbf{1} - v\})$ *such that the correlation structure of P' is ρ.*

Let $\mathcal{B}_n$ denote the set of all n-variate symmetric Bernoulli distributions, E_n the vector containing ordered pairs $\{(i,j) : 1 \le i < j \le n\}$, and let $R : \mathcal{B}_n \to [-1,1]^{E_n}$ map a distribution to its correlation structure. So for a distribution $P \in \mathcal{B}_n$, the correlation vector is

$$R(P) = (\rho_{12}, \rho_{13}, \ldots, \rho_{n-1,n}).$$

The set of all admissible correlation structures is then just $R(\mathcal{B}_n)$.

Let $P_v \sim \mathsf{Unif}(\{v, \mathbf{1} - v\})$ for $v \in \{0,1\}^n$ and $\text{conv}\{P_v : v \in \{0,1\}^n\}$ be the set of all convex combinations of P_v. With this notation, Theorem 1 can be stated as

$$R(\mathcal{B}_n) = R\left(\text{conv}\left\{P_v : v \in \{0,1\}^n\right\}\right).$$

Proof (Proof of Theorem 1) Since each P_v is in $\mathcal{B}_n$, and $\mathcal{B}_n$ is a convex set, we immediately have $R\left(\text{conv}\left\{P_v : v \in \{0,1\}^n\right\}\right) \subseteq R(\mathcal{B}_n)$.

For the other direction, let $P \in \mathcal{B}_n$. So for $X = (X_1, \ldots, X_n) \sim P$, $X_i \sim \mathsf{Bern}(1/2)$ for all i. Note that $X_i \sim 1 - X_i$, so the distribution of $(1 - X_1, \ldots, 1 - X_n)$ is also in $\mathcal{B}_n$ and since $\text{Cor}(X_i, X_j) = \text{Cor}(1 - X_i, 1 - X_j)$ the vector $(1 - X_1, \ldots, 1 - X_n)$ has the same correlation structure as $(X_1, \ldots, X_n)$. Let P^- be the distribution of $(1 - X_1, \ldots, 1 - X_n)$.

Now for any two multivariate symmetric Bernoulli distributions with the same correlation structure, any convex combination of the distributions will have the same covariances, and so the same correlation structure. This convex combination will also still be in $\mathcal{B}_n$. In particular, $P' = (1/2)P + (1/2)P^- \in \mathcal{B}_n$ and $R(P') = R(P)$. For $Y = (Y_1, \ldots, Y_n) \sim P'$ and vector $v \in \{0,1\}^n$,

$$\mathbb{P}(Y = v) = \frac{1}{2}\mathbb{P}(X = v) + \frac{1}{2}\mathbb{P}(X = \mathbf{1} - v) = \mathbb{P}(Y = \mathbf{1} - v).$$

So we can write

$$P' = \sum_{v \in \{0,1\}^n : v(1) = 0} \left[\mathbb{P}(Y = v) + \mathbb{P}(Y = \mathbf{1} - v)\right] P_v,$$

where $P_v \sim \mathsf{Unif}(\{v, \mathbf{1} - v\})$. Hence $P' \in \text{conv}\{P_v : v \in \{0,1\}^n\}$ and since $R(P') = R(P)$ we are done. □

Since the correlation mapping R is affine, the above theorem says that ρ can be a correlation for an n-variate symmetric Bernoulli distribution if and only if it can be written as a convex combination of $R(P_v)$, for $v \in \{0,1\}^n$.

The CUT$_n$ polytope

Related to this is the notion of a cut vector. For a vector $v \in \{0,1\}^n$, let $s(v) = \{i : v_i = 1\}$ be a subset of $[n] = \{1, 2, \ldots, n\}$. Then the partition $\left\{s(v), s(v)^C\right\}$ is a *cut* of K_n, the complete graph with nodes $[n]$.

To any cut can be associated a function on the edges of K_n that will assign 1 to an edge that crosses the cut and 0 otherwise, called cut vector, and this correspondence is one-to-one.

Definition 1 *For every $A \subseteq [n]$ the vector $c^A \in \{0,1\}^{E_n}$ defined as*

$$c^A_{ij} = \begin{cases} 1, \text{ if } |A \cap \{i,j\}| = 1 \\ 0, \quad \text{otherwise} \end{cases}$$

for $(1 \le i < j \le n)$, is called a cut vector of K_n.

For such a cut vector c^A, let $t\left(c^A\right) = A$ if $1 \in A$, otherwise $t\left(c^A\right) = A^C$ $\left(\text{note that } c^A = c^{A^C}\right)$.

Example: take $n = 3$ and $v = (1, 1, 0)$. Then $s(v) = \{1, 2\}$ and the partition $\{\{1, 2\}, \{3\}\}$ is a cut of K_3. Now, for $A = \{1, 2\}$, $A^C = \{3\}$, and $c^A_{12} = 0$, $c^A_{13} = 1$, $c^A_{23} = 1$. Also $t\left(c^{\{1,2\}}\right) = t\left(c^{\{3\}}\right) = \{1, 2\}$.

For a distribution P over $\{0,1\}^n$, let $C(P)$ denote the *concurrence vector*, where if $(X_1, \ldots, X_n) \sim P$, $C(P)(\{i,j\}) = \mathbb{P}(X_i = X_j)$. The set of concurrence vectors are related to the set of cut vectors as follows.

Lemma 1 *Let P be a probability distribution on $\{0,1\}^n$. Then the concurrence vector $C(P)$ is in the convex hull of the set $\{\mathbf{1} - c : c$ is a cut vector of $K_n\}$.*

Proof Let $(X_1, \ldots, X_n) \sim P$. Then

$$\begin{aligned} \mathbb{P}(X_i = X_j) &= \sum_{A:\{i,j\}\subseteq A \text{ or } \{i,j\}\subseteq A^C} \mathbb{P}(s(X) = A) \\ &= \sum_{\text{cut vector } c:c_{ij}=0} \left[\mathbb{P}(s(X) = t(c)) + \mathbb{P}\left(s(X) = t(c)^C\right)\right] \\ &= \sum_{c \text{ a cut vector}} \left[\mathbb{P}(s(X) = t(c)) + \mathbb{P}\left(s(X) = t(c)^C\right)\right](1 - c_{ij}) \end{aligned}$$

Since $\mathbb{P}(s(X) = t(c)) + \mathbb{P}\left(s(X) = t(c)^C\right)$ are nonnegative and sum to 1 over all cut vectors c of K_n, the proof is finished. □

The convex hull of the cut vectors c is known as the CUT$_n$ polytope (see (Deza and Laurent 1997) for details). So another way to state the lemma is that the set of concurrence vectors lies in $\mathbf{1} - \text{CUT}_n$.

For symmetric Bernoullis, the concurrence vector and the correlation structure are directly connected. It is easy to show that $\rho_{ij} := \text{Cor}(X_i, X_j) = 4\mathbb{P}(X_i = X_j = 1) - 1$. Since each $X_i \sim \mathsf{Unif}(\{0,1\})$, $2\mathbb{P}(X_i = X_j = 1) = \mathbb{P}(X_i = X_j)$. Hence $\rho = 2C(P) - \mathbf{1}$, so $(\mathbf{1} + \rho)/2 = C(P) \in \mathbf{1} - \text{CUT}_n$. Finally we have the following.

Theorem 2 *The vector $\rho \in [-1,1]^{E_n}$ is an admissible correlation for the multivariate symmetric Bernoulli family, that is, $\rho \in R(\mathcal{B}_n)$ if and only if $(\mathbf{1}-\rho)/2 \in CUT_n$.*

This result is similar in spirit to work of Avis (1977), and in fact can also be derived from his results.

Simulation from multivariate distributions with given correlations

In general, creating a multivariate symmetric Bernoulli distribution with specified correlations can be done by testing feasibility of a linear program. The program contains 2^n decision variables, one for each $v \in \{0,1\}^n$, and x_v represents the probability that $X = v$. There is one equality constraint for each $i \in \{1,\ldots,n\}$:

$$\sum_{v:v(i)=1} x_v = 1/2.$$

There are $\binom{n}{2}$ equality constraints for each of the correlations:

$$\sum_{v:v(i)=v(j)} x_v - \sum_{v:v(i)\neq v(j)} x_v = \rho_{ij},$$

and a final equality constraint

$$\sum_v x_v = 1.$$

Last, the x_v must be nonnegative.

By employing Theorem 1, we can cut the number of decision variables in the linear program in half, since each diagonal of $[0,1]^n$ is described by a vector $v \in \{0,1\}^n$ with $v(1) = 0$. Let α_v denote these decision variables. Then because we are mixing uniforms over $\{v, \mathbf{1}-v\}$, the $\sum_{v:v(i)=1} x_v = 1/2$ constraints are automatically satisfied. All that remain are the correlation, total sum, and nonnegativity constraints.

$$(\forall i,j)\left(\sum_{v:v(i)=v(j)} \alpha_v - \sum_{v:v(i)\neq v(j)} \alpha_v = \rho_{ij}\right), \ \sum_v \alpha_v = 1, \text{ and } (\forall v)(\alpha_v \geq 0).$$

To illustrate this procedure, suppose that we wish to simulate draws from (T_1, T_2, T_3) where the T_i are exponential random variables with rate 1 and correlation structure

$$\text{Cor}(T_1,T_2) = 0.7,\ \text{Cor}(T_1,T_3) = -0.4, \text{Cor}(T_2,T_3) = -0.2.$$

The following procedure is given in Huber and Marić (2015). Recall that for $U \sim \text{Unif}([0,1])$, the inverse transform method gives that both $-\ln(U)$ and $-\ln(1-U)$ have an exponential distribution with rate 1.

Suppose that $\text{Cor}(B_1,B_2) = 0.635244$. Then draw $U \sim \text{Unif}([0,1])$, and let $T_i = -\ln(U)B_i + -\ln(1-U)(1-B_i)$. Then it is an easy calculation to show that $\text{Cor}(T_1,T_2) = 0.7$. Similarly, by generating

$$\text{Cor}(B_1,B_2) = 0.635244,\ \text{Cor}(B_1,B_3) = -0.70220,\ \text{Cor}(B_2,B_3) = -0.45903.$$

and calculating the T_i in the same fashion, the complete correlation structure for (T_1, T_2, T_3) can be replicated.

Because for symmetric Bernoullis $\text{Cor}(B_i,B_j) = 4\text{Cov}(B_i,B_j)$ and covariance is an inner product, the correlation of a convex combination of variables is the convex combination of the correlations. By the symmetry of $\{0,1\}^n$, we need only consider vectors with first component 0. Hence the vectors to consider are $(v_1, v_2, v_3, v_4) = ((0,0,0), (0,0,1), (0,1,0), (0,1,1))$. For a draw from the distribution where $\mathsf{Unif}(\{v_i, 1-v_i\})$ has coefficient α_i, the correlations would be

$$\alpha_1 + \alpha_2 - \alpha_3 - \alpha_4 = \mathrm{Cor}(B_1, B_2) = 0.635244$$
$$\alpha_1 - \alpha_2 + \alpha_3 - \alpha_4 = \mathrm{Cor}(B_1, B_3) = -0.70220$$
$$\alpha_1 - \alpha_2 - \alpha_3 + \alpha_4 = \mathrm{Cor}(B_2, B_3) = -0.45903$$

Finally, $\sum_i \alpha_i = 1$.

In general, to determine if these equations have a solution we would determine feasibility of a linear program with the additional nonnegativity constraint that all $\alpha_i \geq 0$. In this case, since $\binom{3}{2} + 1 = 2^{3-1}$ there is but one unique solution:

$$(\alpha_1, \alpha_2, \alpha_3, \alpha_4) = (0.1185035, 0.6991185, 0.0303965, 0.1519815).$$

Since these all lie in $[0,1]$, these correlations are admissible.

Our procedure then is to draw a random variable N using $\mathbb{P}(N = i) = \alpha_i$. Next draw $U \sim \mathsf{Unif}([0,1])$. If the i-th component of v_N is 1, then $T_i = -\ln(U)$. Otherwise $T_i = -\ln(1-U)$. As shown in Huber and Marić (2015), this creates a vector (T_1, T_2, T_3) with the desired marginals.

Discussion

Characterizing $R(\mathcal{B}_n)$ via its extreme points naturally raises the same question about the convex set $\mathcal{B}_n$. Even though clearly every P_v is an extreme point of $\mathcal{B}_n$, it should be noted that $\mathcal{B}_n \neq conv\{P_v : v \in \{0,1\}^n\}$. Gérard Letac (private communication) gives an example in $n = 3$ that confirms this statement: a measure that assigns weight 1/4 to $(1,1,1), (1,0,0), (0,1,0), (0,0,1)$ is not a convex combination of P_v's but it clearly belongs to $\mathcal{B}_3$ and moreover is also an extreme point of that set. Characterization of $\mathcal{B}_n$ is still an open problem.

It should be noted that the relation between CUT_n and $\mathcal{B}_n$ does not extend to asymmetric multivariate Bernoulli distributions. It is enough to analyze the bivariate case with equal marginals. The correlation between two $\mathsf{Bern}(p)$ random variables belongs to the interval $[\rho_{\min}, 1]$. Maximum correlation in case of equal marginals, always equals to 1 and the minimum correlation $\rho_{\min}$ can be calculated using Fréchet-Hoeffding bounds (Fréchet 1951; Hoeffding 1940)

$$\rho_{\min} = \begin{cases} -(1-p)/p, & \text{for } p \geq 1/2 \\ -p/(1-p), & \text{for } p \leq 1/2. \end{cases}$$

It is clear now that only for $p = 1/2$, $\rho_{\min} = -1$ and possible correlations equal to the entire interval $[-1,1]$, while for any other value of p it is a strict subinterval of $[-1,1]$. For example, for $p = 3/4$, $-1/3 \leq \rho \leq 1$.

In two dimensional case the cut polytope is known to be $\mathrm{CUT}_2 = [0,1]$ so it corresponds to $R(\mathcal{B}_2)$ only in the symmetric case.

It should be noted also a relation with the elliptope $\mathcal{E}_n$. The set of $n \times n$ correlation matrices is a nonpolyhedral convex set with a nonsmooth boundary and its extreme points of have not been explicitly determined, but there exist characterization results on the rank one and two extreme points, done by Ycart (1985) (see also Li and Tam (1994) and Parthasarathy (2002)). Laurent and Poljak (1995) proved that cut matrices (analogous to cut vectors) are actually vertices-rank one extreme point of the elliptope and that $\mathcal{E}_n$ can be seen as a nonpolyhedral relaxation of the cut polytope. In view of theorems proved here it follows that the vertices of $\mathcal{E}_n$ correspond precisely to symmetric Bernoulli correlations.

Abbreviations
$\mathcal{B}_n$: Set of all n-variate symmetric Bernoulli distributions; Bern(1/2): Symmetric Bernoulli distribution; $conv\{S\}$: Convex hull of a finite point set S i.e. the set of all convex combinations of its points; Cor(X, Y): Correlation between random variables X and Y; Cov(X, Y): Covariance between random variables X and Y; CUT$_n$: Convex hull of cut vectors in a complete graph with vertices $\{1, \ldots, n\}$; $\mathcal{E}_n$: (The elliptope) set of all symmetric positive semi-definite $n \times n$ matrices that have all ones on the diagonal; ρ_{min}: Minimum possible correlation among two random Bernoulli variables; Unif(S): Uniform distribution over finite set S

Acknowledgements
The authors are grateful to Gérard Letac for sharing his ideas with them and for inspiring discussions.

Funding
MH was partially supported by NSF grant DMS-1418495. NM was partially supported by a University of Missouri Research Board award.

Availability of data and materials
Not applicable.

Authors' contributions
The authors MH and NM carried out this work and drafted the manuscript together. Both authors read and approved the final manuscript.

Competing interests
The authors declare that they have no competing interests.

Publisher's Note

Author details
[1]Claremont McKenna College, 850 Columbia Av, Claremont 91711-3901, CA, USA. [2]University of Missouri-St. Louis, 1 University Blvd, 311 Express Scripts Hall, St. Louis 63121-4400, MO, USA.

Received: 17 September 2018 Accepted: 4 February 2019
Published online: 08 March 2019

References
Avis, D: Some Polyhedral Cones Related to Metric Spaces. Ph. D. Thesis, Stanford University (1977)
Chaganty, NR, Joe, H: Range of correlation matrices for dependent Bernoulli random variables. Biometrika. **93**(1), 197–206 (2006)
Devroye, L, Letac, G: Copulas with Prescribed Correlation Matrix. In: Memoriam Marc Yor-Séminaire de Probabilités XLVII, pp. 585–601. Springer, Cham, (2015)
Deza, MM, Laurent, M: Geometry of Cuts and Metrics. Algoritm Combin. **15** (1997)
de Finetti, B: A proposito di correlazione. Supplemento Statistico ai Nuovi problemi di Politica Storia ed Economia. **3**, 41–57 (1937)
dos Santos Dias, CT, Samaranayaka, A, Manly, B: On the use of correlated beta random variables with animal population modelling. Ecol Model. **215**(4), 293–300 (2008)
Fréchet, M: Sur les tableaux de corrélation dont les marges sont données. Ann. Univ. Lyon, 3^e serie, Sciences, Sect. A. **14**, 53–77 (1951)
Henderson, SG, Chiera, BA, Cooke, RM: Generating dependent quasi-random numbers. In: Proceedings of the 32nd conference on Winter simulation, pp. 527–536, (2000). Society for Computer Simulation International
Hill, RR, Reilly, CH: Composition for multivariate random variables. In: Tew, J, Manivannan, S, Sadowski, D, Seila, A (eds.) Proceedings of the 1994 Winter Simulation Conference, pp. 332–339, (1994). http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=717172
Hoeffding, W: Massstabinvariante Korrelationstheorie. Schriften des Mathematischen Instituts und des Instituts für Angewandte Mathematik der Universitat Berlin. **5**, 179–233 (1940)
Huber, M, Marić, N: Simulation of multivariate distributions with fixed marginals and correlations. J Appl Probab. **52**(2), 602–608 (2015). arXiv:1311.2002
Lampard, DG: A stochastic process whose successive intervals between events form a first-order Markov chain. J Appl Probab. **5**, 648–668 (1968)
Laurent, M, Poljak, S: On a positive semidefinite relaxation of the cut polytope. Linear Algebra Appl. **223**, 439–461 (1995)

Lawrance, AJ, Lewis, PAW: A new autoregressive time series model in exponential variables (NEAR). Adv Appl Probab. **13**(4), 826–845 (1981)
Leonov, S, Qaqish B: Correlated endpoints: simulation, modeling, and extreme correlations. Statist. Papers To appear
Lévy, P: Distance de deux variables aléatoires et distance de deux lois de probabilité. Traité, de calcul des probabilités et de ses applications by Emile Borel. **I**(III), 286–292 (1937)
Li, C-K, Tam, B-S: A note on extreme correlation matrices. SIAM J Matrix Anal Appl. **15**(3), 903–908 (1994)
Parthasarathy, KR: On extremal correlations. J Stat Plan Infer. **103**(1), 173–180 (2002)
Smith, OE, Adelfang, SI: Gust model based on the bivariate gamma probability distribution. J Spacecr Rocket. **18**, 545–549 (1981)
Ycart, B: Extreme points in convex sets of symmetric matrices. Proc Am Math Soc. **95**(4), 607–612 (1985)