Journal of Statistical Modelling: Theory and Applications
Vol. 5, No. 2, 2024, pp. 117-130
Yazd University Press 2024

JSMTA

Research Paper

Optimal design for a Fréchet copula marginal regression with exponential marginals

Reza Farhadian and Habib Jafari*
Department of Statistics, Razi University, Kermanshah, Iran

Received: July 19, 2024/ Revised: October 12, 2024/ Accepted: November 28, 2024

Abstract: Modeling dependence using copula functions is a standard and widely used method in applied statistics. In recent years, experiments have been conducted in which several dependent responses are described by regression models, and then experimental designs for these models have been carried out. These regression models are often expressed as copula-based regression models, so the copula parameters (if they exist) also affect the optimal design problem. In this paper, we consider the dependence structure of a random pair from the exponential distribution conditionally upon a covariate as a regression model, then investigate the D-optimal design for this copula-based regression model. The copula function that is used is the Fréchet copula. The optimal designs obtained all have a general form depending on the selected design space.

Keywords: Copula-based model; D-optimal design; Fisher information matrix; Fréchet copula.
Mathematics Subject Classification (2010): 62K05, 62H20.

1 Introduction

Optimal design in statistical modeling has been studied and developed for a century. Over 100 years ago, Smith proposed (Smith, 1918) a criterion based on which she obtained optimal designs for regression problems. Smith showed that in order to minimize the variance of the least squares estimate of the slope in a simple linear regression model, one should take observations in equal numbers at the two extremes of the range

*Corresponding author: `h.jafari@razi.ac.ir`

of the explanatory variable. Many years later, Kiefer (1959) developed useful computational methods for finding optimal designs in linear regression issues (see also Kiefer and Wolfowitz, 1959).

After these pioneering works, the topic of optimal design was widely developed, and numerous studies have been conducted in this field, so that nowadays studying and modeling data without using optimal designs is practically not cost-effective. In the context of experimental design, the concept of optimal design refers to a specific category of experimental design that is classified based on certain statistical criteria. Usually, in model-based optimal designs, the inferential aim is to estimate the parameters of the model, so that estimators with minimum variance are of interest. The optimization process is completely dependent on the considered optimality criteria, which are usually defined in terms of the information matrix. In linear models, the information matrix and then optimality criteria do not depend on the unknown parameters of the model, therefore, reaching an optimal design does not have significant computational complexity. However, in non-linear models, since unknown parameters often appear in the entries of the information matrix, the optimality criteria depend on these parameters, so deriving an optimal design from the optimization problem may involve computational complexities (see, e.g, Atkinson et al., 2007; Burkner et al., 2019). Therefore, in relation to the optimal design in non-linear models, the primary challenge is to remove unknown parameters from the optimality criteria. Various methods and solutions have been proposed to deal with this challenge, some of the most typical of which are local optimal design, sequential optimal design, minimax optimal design, Bayesian optimal design, and pseudo-Bayesian optimal design. For some references on these methods, see, e.g., Chernoff (1953), Kiefer and Wolfowitz (1959), Kiefer (1974), Dette et al. (2006), Graßhoff et al. (2012), Parsa Maram and Jafari (2016), and Aminnejad and Jafari (2017).

In local optimality, certain values are selected (based on prior studies or experiments) as the best guess for the unknown parameters of the model, and then a function of the information matrix is optimized to derive designs for the values of these parameters. Since the guess values may not always express the behavior of the parameters well, a prior distribution is used instead of guess values; this method is known as Bayesian optimality, and the design derived from this method is called the Bayesian optimal design. The Bayesian optimal design is based on maximizing the expected utility function for the model. A prior distribution for the parameters is chosen based on prior information and the researcher's beliefs. However, there is no definitive method for selecting the best prior distribution. Numerous researchers have investigated the effect of the prior distribution on determining design points in various types of optimal designs. For instance, Chaloner and Duncan (1983), Mukhopadhyay and Haines (1995), Dette and Neugebauer (1997), Firth and Hinde (1997), Fedorov and Hackl (2012), Burghaus and Dette (2014), Goudarzi et al. (2019), and Abdollahi et al. (2024). For a more advanced study on this topic, see Chapter 18 of the book by Atkinson et al. (2007).

In recent years, the experimental design has been used in dependency modeling through copula-based models. The concept of copula in experimental design is new and has been around for a little more than a decade, considering works such as Li et al. (2011), Pilz et al. (2012), and Schmidt et al. (2014). The design of experiments for copula parameter estimation was first proposed in Denman et al. (2011), where a brute-force simulated annealing optimization was employed for the solution of a specific problem. After that, extensive studies were conducted. Here, we briefly address some of these recent developments. Perrone and Müller (2016) presented an equivalence theorem for bivariate copula models, which allows for the formulation of efficient design algorithms and a quick check of whether designs are optimal or at least efficient. That comprehensive work included some examples and comparisons between different copula models with respect to design efficiency. These methods were extended to the local D_A-criterion (and, as a special case, to the D_s-criterion) in the work of Perrone et al. (2016), where a wide range of flexible copula models were analyzed to highlight the usefulness of D_s-optimality in many possible scenarios. Durante and Perrone (2016) provided an overview of the definitions and properties of asymmetric copulas, which are copulas whose values are not constant under any change in their arguments, then discussed how asymmetric copulas may also be useful in the optimal design of experiments. Deldossi et al. (2018) considered a bivariate logistic model for a binary response and, assuming two possible rival dependence structures, used copula functions for modeling different kinds of dependence with arbitrary marginal distributions. Rappold et al. (2020) developed designs with blocks of size two using copula models and adopted a pseudo-Bayesian approach to constructing block designs.

These studies on optimal design for copula models motivated us to develop this idea for other copula-based regression models. This research work deals with a copula-based regression model with exponentially distributed marginals, where the model is influenced by only one covariate point. The copula function that is used is the Fréchet copula, which is a well-known parametric copula. Afterward, the optimal design for these models is derived under the D-optimal criterion.

The paper is organized as follows. Section 2 presents a theoretical framework based on copula theory and optimal design theory. The Fréchet copula and some of its properties are briefly given in Section 3. In Section 4, the optimal design for a copula-based model with exponential marginals is derived. Section 5 concludes this paper.

2 Optimal experimental design for copula models

2.1 Copula: a tool for modeling dependence

Let Y_1 and Y_2 be arbitrary random variables with a joint distribution function $F_{Y_1,Y_2}(y_1, y_2)$ and marginals $F_{Y_1}(y_1)$ and $F_{Y_2}(y_2)$, respectively. When Y_1 and Y_2 are independent, then we can write $F_{Y_1,Y_2}(y_1, y_2) = F_{Y_1}(y_1)F_{Y_2}(y_2)$. In the case of dependency, this equality is replaced by $\geq$ or $\leq$, and the direction of these inequalities may not hold for all y_1, y_2. However, if $F_{Y_1,Y_2}(y_1, y_2) \geq F_{Y_1}(y_1)F_{Y_2}(y_2)$ for all y_1, y_2, then Y_1 and Y_2 are said to be positively dependent, and if $F_{Y_1,Y_2}(y_1, y_2) \leq F_{Y_1}(y_1)F_{Y_2}(y_2)$ for all y_1, y_2, then Y_1 and Y_2 are said to be negatively dependent (see, e.g. Lehmann, 1966; Joag-Dev and Proschan, 1983). Therefore, in dependency cases, to specify the joint distribution $F_{Y_1,Y_2}(y_1, y_2)$ in terms of marginal distributions $F_{Y_1}(y_1)$ and $F_{Y_2}(y_2)$, a function is used to combine the marginal distributions to construct the joint distribution. This associative function is called the *copula function.* Note that a copula is k-dimensional if the distribution function $F(.)$ is k-dimensional. Here we focus on $k = 2$, which is called a bivariate copula, a 2-dimensional copula, or, more briefly, a 2-copula. Following

Nelsen (2006), a 2-copula is defined as follows

Definition 2.1. *A 2-copula is a bivariate function* $C(u, v) : [0, 1] \times [0, 1] \to [0, 1]$ *which holds under the following conditions:*
(I) for every $u, v \in [0, 1]$,

$$C(u, 0) = 0, \quad C(u, 1) = u, \quad C(0, v) = 0, \quad C(1, v) = v,$$

(II) for every $u_1, v_1, u_2, v_2 \in [0, 1]$ *with* $u_1 \leq v_1$ *and* $u_2 \leq v_2$,

$$C(u_2, v_2) - C(u_2, v_1) - C(u_1, v_2) + C(u_1, v_1) \geq 0.$$

According to Definition 2.1, it can be concluded that the copula $C(u, v)$ is a bivariate distribution function with uniform marginals on $[0, 1]$. By Sklar's Theorem (see Sklar, 1959), this copula exists for every joint distribution function $F_{Y_1,Y_2}(y_1, y_2)$ with marginal distributions $F_{Y_1}(y_1)$ and $F_{Y_2}(y_2)$. Moreover, the copula $C(u, v)$ is only unique on $\text{Ran}(F_{Y_1}) \times \text{Ran}(F_{Y_2})$, where $\text{Ran}(F_{Y_i})$ $(i = 1, 2)$ denotes the range of the distribution function F_{Y_i}. Thus, we can construct a bivariate distribution function $F_{Y_1,Y_2}(y_1, y_2)$ as

$$F_{Y_1,Y_2}(y_1, y_2) = C(F_{Y_1}(y_1), F_{Y_2}(y_2)), \qquad \text{for all } y_1, y_2.$$

The partial derivatives $\frac{\partial C(u,v)}{\partial u}$ and $\frac{\partial C(u,v)}{\partial v}$ exist, and $c(u, v) = \frac{\partial^2 C(u,v)}{\partial u \partial v}$ is the probability density function of $C(u, v)$. Hence, if $F_{Y_1}(y_1)$ and $F_{Y_2}(y_2)$ are continuous with densities $f_{Y_1}(y_1)$ and $f_{Y_2}(y_2)$, then the corresponding joint density function $f_{Y_1,Y_2}(y_1, y_2)$ is obtained through equation $f_{Y_1,Y_2}(y_1, y_2) = \frac{\partial^2 C(F_{Y_1}(y_1), F_{Y_2}(y_2))}{\partial y_1 \partial y_2} = c\left(F_{Y_1}(y_1), F_{Y_2}(y_2)\right) \times f_{Y_1}(y_1) \times f_{Y_2}(y_2)$.

An important (and simple) copula is the product copula $C(u, v) = \prod(u, v) = uv$ (with $0 \leq u, v \leq 1$), which is corresponds to independence. Except for this simple one, copulas usually have one or more parameters in their functional structure. Thus, every parametric copula is denoted by $C(.; \boldsymbol{\alpha})$, where $\boldsymbol{\alpha} = (\alpha_1, \alpha_2, \ldots, \alpha_z)^T$ is a vector of real-valued parameters. Therefore, in addition to selecting the appropriate copula for describing the model, one should be as careful as possible in selecting the copula parameters. There are criteria that are useful in selecting appropriate copula parameters. One of these useful tools is Kendall's tau concordance measure, which is defined as follows

$$\tau = 4 \int_{[0,1]^2} C(u, v; \boldsymbol{\alpha}) dC(u, v; \boldsymbol{\alpha}) - 1 = 4 \int_0^1 \int_0^1 C(u, v; \boldsymbol{\alpha}) dC(u, v; \boldsymbol{\alpha}) - 1,$$

that is, $\tau = 4E[C(U, V)] - 1$, where U and V have support in $[0, 1]$ with joint distribution function $C(u, v; \boldsymbol{\alpha})$. Moreover, if $C(.; \boldsymbol{\alpha})$ is absolutely continuous, then $\tau = 4 \int_{[0,1]^2} C(u, v; \boldsymbol{\alpha}) c(u, v; \boldsymbol{\alpha}) du dv - 1$. For a generalized version of Kendall's tau, see Joe (1990). For more details, examples, and applications of copulas, we refer to Nelsen (2006), Durante and Sempi (2015), and Cherubini et al. (2016).

2.2 Experimental design issue

Let us consider a regression model consisting of a vector $\boldsymbol{x} = (x_1, \ldots, x_d)^T$ of explanatory variables, and a vector $\boldsymbol{y}(\boldsymbol{x}) = (y_1(\boldsymbol{x}), \ldots, y_m(\boldsymbol{x}))^T$ of response variables. Suppose that in a regression experiment, the result of the expectations is the following vector

$$\boldsymbol{E}[\boldsymbol{Y}(\boldsymbol{x})] = \boldsymbol{E}[(Y_1, \ldots, Y_m)^T] = (\eta_1(\boldsymbol{x};\boldsymbol{\theta}), \ldots, \eta_m(\boldsymbol{x};\boldsymbol{\theta}))^T = \boldsymbol{\eta}(\boldsymbol{x}, \boldsymbol{\theta}),$$

where $\boldsymbol{\theta} = (\theta_1, \ldots, \theta_k)^T$ is a vector of unknown regression parameters and $\eta_i(\boldsymbol{x};\boldsymbol{\theta})$'s $(i = 1, 2, \ldots, m)$ are known functions. If $F_{Y_i}(y_i(\boldsymbol{x}, \boldsymbol{\theta}))$ is the marginal distribution function of each Y_i for $i = 1, 2, \ldots, m$, then, based on the Sklar's theorem Sklar (1959), the dependence between Y_i's is modeled by a copula function such as

$$C(\boldsymbol{F_Y}(\boldsymbol{y}(\boldsymbol{x}, \boldsymbol{\theta})); \boldsymbol{\alpha}) = C(F_{Y_1}(y_1(\boldsymbol{x}, \boldsymbol{\theta})), \ldots, F_{Y_m}(y_m(\boldsymbol{x}, \boldsymbol{\theta}); \boldsymbol{\alpha})),$$

where $\boldsymbol{\alpha} = (\alpha_1, \ldots, \alpha_z)^T$ are unknown copula parameters.

Thus, $f_{\boldsymbol{Y}}(\boldsymbol{y}(\boldsymbol{x}, \boldsymbol{\theta}); \boldsymbol{\alpha}) = \frac{\partial^m C(\boldsymbol{F_Y}(\boldsymbol{y}(\boldsymbol{x},\boldsymbol{\theta}));\boldsymbol{\alpha})}{\partial y_1 \partial y_2 \ldots \partial y_m}$ is the joint probability density function of the random vector $\boldsymbol{Y}$. Note that the joint probability mass function is considered in the discrete case. The corresponding Fisher information matrix is a square matrix of order $(k + z)$, which is defined as follows

$$M(\boldsymbol{x}, \boldsymbol{\theta}, \boldsymbol{\alpha}) = \begin{pmatrix} m_{\boldsymbol{\theta}\boldsymbol{\theta}}(\boldsymbol{x}) & m_{\boldsymbol{\theta}\boldsymbol{\alpha}}(\boldsymbol{x}) \\ m^T_{\boldsymbol{\theta}\boldsymbol{\alpha}}(\boldsymbol{x}) & m_{\boldsymbol{\alpha}\boldsymbol{\alpha}}(\boldsymbol{x}) \end{pmatrix},$$

where the submatrices $m_{\boldsymbol{\theta}\boldsymbol{\theta}}(\boldsymbol{x})$ and $m_{\boldsymbol{\alpha}\boldsymbol{\alpha}}(\boldsymbol{x})$ are square matrices of orders $k \times k$ and $z \times z$, respectively, and the submatrix $m_{\boldsymbol{\theta}\boldsymbol{\alpha}}(\boldsymbol{x})$ is of order $k \times z$. The elements of these submatrices are defined as the Fisher information of the probability density function $f_{\boldsymbol{Y}}(\boldsymbol{y}(\boldsymbol{x}, \boldsymbol{\theta}); \boldsymbol{\alpha})$ about the parameters in the subscript of each submatrix. For example, the (i, j)th element of the submatrix $m_{\boldsymbol{\theta}\boldsymbol{\theta}}(\boldsymbol{x})$ is defined as

$$\boldsymbol{E}\left(-\frac{\partial^2}{\partial\theta_i \partial\theta_j} \log f_{\boldsymbol{Y}}(\boldsymbol{y}(\boldsymbol{x}, \boldsymbol{\theta}); \boldsymbol{\alpha})\right), \tag{1}$$

and other two submatrices $m_{\boldsymbol{\alpha}\boldsymbol{\alpha}}(\boldsymbol{x})$ and $m_{\boldsymbol{\theta}\boldsymbol{\alpha}}(\boldsymbol{x})$ are defined accordingly.

Now, the aim of optimal design theory is to quantify the amount of information on parameters $\boldsymbol{\theta}$ and $\boldsymbol{\alpha}$ from the regression experiment embodied in the Fisher information matrix $M(\boldsymbol{x}, \boldsymbol{\theta}, \boldsymbol{\alpha})$. An approximate design ξ for this model is a probability measure on the design space $\boldsymbol{\chi}$ with finite support $x_1, x_2, \ldots, x_r$ and weights $w_1, w_2, \ldots, w_r$ assigned to x_i's. Usually, the design ξ is denoted by

$$\xi = \begin{Bmatrix} x_1 & x_2 & \ldots & x_r \\ w_1 & w_2 & \ldots & w_r \end{Bmatrix} \in \boldsymbol{\Xi}, \tag{2}$$

where $\boldsymbol{\Xi} = \{\xi | 0 \le w_i \le 1; \sum_{i=1}^r w_i = 1, x \in \boldsymbol{\chi}\}$ is the space of all possible designs (see, e.g. Kiefer, 1974)).

Note that the number of points in the optimal design satisfies in the inequality $p \le r \le \frac{p(p+1)}{2}$, where p is the number of parameters (Silvey (1980)). For the model described above, we have $p = k + z$. The information matrix of the design in (2) is defined as follows (see, e.g., Atkinson et al., 2007)

$$\mathcal{M}(\xi, \boldsymbol{\theta}, \boldsymbol{\alpha}) = \sum_{i=1}^r w_i M(x_i, \boldsymbol{\theta}, \boldsymbol{\alpha}). \tag{3}$$

The optimal design problem is finding $\xi^* = \left\{ \begin{smallmatrix} x_1^* & x_2^* & \cdots & x_r^* \\ w_1^* & w_2^* & \cdots & w_r^* \end{smallmatrix} \right\}$, which maximizes a function $\psi(.)$ (which is defined by a proper criterion) of the information matrix $\mathcal{M}(\xi, \boldsymbol{\theta}, \boldsymbol{\alpha})$, or in other words, $\xi^* = \arg\max_{\xi \in \boldsymbol{\Xi}} \psi\big(\mathcal{M}(\xi, \boldsymbol{\theta}, \boldsymbol{\alpha})\big)$ (see, for example, Pukelsheim, 1993). One of the most common such criteria is *D-optimality*, that is, the criterion $\psi(\mathcal{M}(\xi, \boldsymbol{\theta}, \boldsymbol{\alpha})) = \log(\det(\mathcal{M}(\xi, \boldsymbol{\theta}, \boldsymbol{\alpha})))$ for when $\mathcal{M}(\xi, \boldsymbol{\theta}, \boldsymbol{\alpha})$ is non-singular (Silvey, 1980). Moreover, when the elements obtained through (1) include unknown parameters $\boldsymbol{\theta}$ and $\boldsymbol{\alpha}$, obviously the information matrix $\mathcal{M}(\xi, \boldsymbol{\theta}, \boldsymbol{\alpha})$ also depends on the unknown parameters. Therefore, to implement a design, the role of the parameter in the information matrix must be solved first. A typical traditional method is the local optimal method, which is methodologically simple. However, the efficiency of the optimal designs derived from this method depends on the guess values to substitute for the unknown parameters. For instance, in locally optimization, the D-optimal design is defined as $\xi^* = \arg\max_{\xi \in \boldsymbol{\Xi}} \log\big(\det\big(\mathcal{M}(\xi, \bar{\boldsymbol{\theta}}, \bar{\boldsymbol{\alpha}})\big)\big)$, where $(\bar{\boldsymbol{\theta}}, \bar{\boldsymbol{\alpha}})$ is the local guesses (fixed values) of the unknown parameter vector $(\boldsymbol{\theta}, \boldsymbol{\alpha})$. A more robust approach is to use a prior distribution instead of guess values for the unknown parameters, which is the concept of the Bayesian optimal design. In this regard, the Bayesian *D-optimality* criterion is one of the most common criteria. A design is called Bayesian D-optimal with respect to a given prior π on $\boldsymbol{\theta}$ if it maximizes the function

$$\Psi_\pi(\xi) = E[\psi(\mathcal{M}(\xi, \boldsymbol{\theta}, \boldsymbol{\alpha}))] = \int_{\boldsymbol{\theta}} \psi(\mathcal{M}(\xi, \boldsymbol{\theta}, \boldsymbol{\alpha}))\pi(d\boldsymbol{\theta}), \tag{4}$$

where $\psi(\mathcal{M}(\xi, \boldsymbol{\theta}, \boldsymbol{\alpha})) = \log(\det(\mathcal{M}(\xi, \boldsymbol{\theta}, \boldsymbol{\alpha})))$ is the D-optimality criterion. Therefore, the Bayesian D-optimal design is defined as $\xi^* = \arg\max_{\xi \in \boldsymbol{\Xi}} \Psi_\pi(\xi)$.

There are theorems that are useful in finding and checking D-optimal designs. One of the most famous of these theorems is the Kiefer-Wolfowitz–type equivalence theorem (Kiefer and Wolfowitz, 1960; Silvey, 1980). Perrone and Müller (2016) provided an equivalence theorem for 2-copula models which states that for a local parameter vector $(\bar{\boldsymbol{\theta}}, \bar{\boldsymbol{\alpha}})$ (local guesses), the following properties are equivalent: (i) ξ^* is D-optimal; (ii) $d(x, \xi^*) \le (k+1)$ for $x \in \boldsymbol{\chi}$; (iii) ξ^* minimize $\max d(x, \xi^*)$ over all $\xi \in \boldsymbol{\chi}$. Note that $d(x, \xi^*) = trace\big(\mathcal{M}(\xi^*, \bar{\boldsymbol{\theta}}, \bar{\boldsymbol{\alpha}})^{-1} M(x, \bar{\boldsymbol{\theta}}, \bar{\boldsymbol{\alpha}})\big)$, which is usually called *sensitivity function.* See also the work of Perrone et al. (2016) for a similar equivalence theorem for so-called D_A and D_s criteria for copula models. There are also methods to compare designs, one of the most important of which is a ratio called *efficiency.* In particular, if the optimization problem is performed under the locally D-optimality, then the D-efficiency of the design ξ with respect to the D-optimal design ξ^* is the ratio $D_{eff} = \left(\frac{\det(\mathcal{M}(\xi, \bar{\boldsymbol{\theta}}, \bar{\boldsymbol{\alpha}}))}{\det(\mathcal{M}(\xi^*, \bar{\boldsymbol{\theta}}, \bar{\boldsymbol{\alpha}}))}\right)^{\frac{1}{p}}$, where p is the number of the model parameters and $(\bar{\boldsymbol{\theta}}, \bar{\boldsymbol{\alpha}})$ is the local guesses of the parameter vector $(\boldsymbol{\theta}, \boldsymbol{\alpha})$.

3 The Fréchet copula

Some parametric copulas are constructed based on the fact that any convex combination of arbitrary copulas perserves the properties of a copula function (stated in Definition 2.1). A well-known family of such copulas is the Fréchet family. A bivariate Fréchet copula is obtained through a convex combination of the product cupula

$\prod(u,v)$, the comonotonicity copula $\mathbb{M}(u,v) = \min(u,v)$, and the countermonotonicity copula $\mathbb{W}(u,v) = \max(u+v-1,0)$ as follows (Xie et al., 2022)

$$C^{\boldsymbol{Fre}}(u,v;\alpha,\beta,\gamma) = \alpha\mathbb{M}(u,v) + \gamma\prod(u,v) + \beta\mathbb{W}(u,v), \tag{5}$$

where $\alpha,\beta,\gamma \geq 0$ (with $\alpha+\beta+\gamma=1$) represent the dependence parameters. Note that we can always consider $\gamma = 1-\alpha-\beta$.

Figure 1 shows the surface of the Fréchet copula in (5) for different values of parameters α,β, and γ. The Fréchet copula density is $c^{\boldsymbol{Fre}}(u,v) = \gamma$ (Cherubini et al., 2004). Furthermore, the Kendall's τ associated with the Fréchet copula in (5) is $\tau = \frac{1}{3}(\beta-\alpha)(\alpha+\beta+2)$ (Durante and Sempi, 2015). Note that sometimes the bivariate copulas $\mathbb{M}(u,v)$ and $\mathbb{W}(u,v)$ are called respectively the Fréchet upper bound and the Fréchet lower bound, and it is well-known that for any bivariate copula $C(u,v)$, we always have $\mathbb{W}(u,v) \leq C(u,v) \leq \mathbb{M}(u,v)$. Therefore, the Fréchet copula in (5), which characterizes dependence as a mixture of three simple structures of independence, comonotonicity, and countermonotonicity, is useful for approximating any bivariate copula, see Salvadori et al. (2007). Moreover, note that it is not possible to generalize the bivariate Fréchet copula to the multivariate case because the countermonotonicity copula $\mathbb{W}(u,v)$ can not be defined as a copula for more than two variables. However, a possible d-dimensional extension of the Fréchet copula in (5) (for $d \geq 2$) can be obtained through a convex combination of the d-dimensional product cupula $\prod_d(\boldsymbol{u}) = u_1u_2\cdots u_d$ and the d-dimensional comonotonic copula $\mathbb{M}_d(\boldsymbol{u}) = \min(u_1,u_2,\ldots,u_n)$, that is, $C^{\boldsymbol{Fre}}(\boldsymbol{u};\alpha) = \alpha\mathbb{M}_d(\boldsymbol{u}) + (1-\alpha)\prod_d(\boldsymbol{u})$, $\boldsymbol{u} \in [0,1]^d$ and $\alpha \in [0,1]$ (Durante and Sempi, 2015).

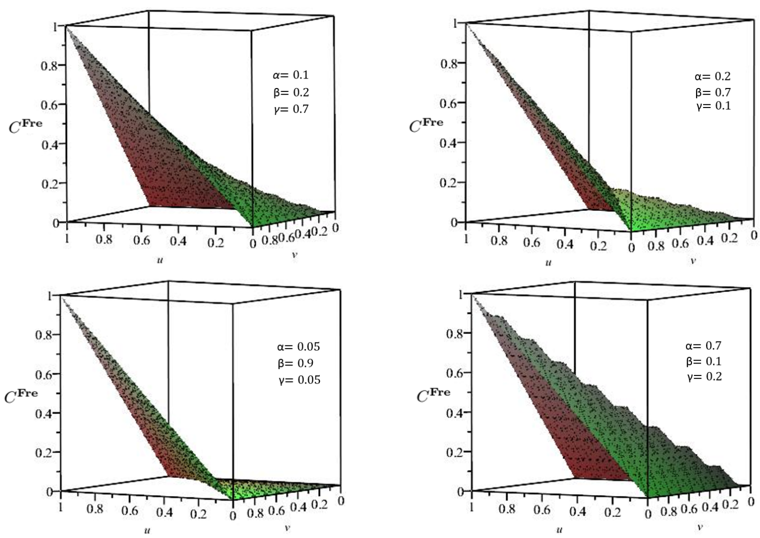

Figure 1: Plots of wireframe surface for the Fréchet copula in (5) for different values of parameters α,β, and γ.

4 Optimal design with copula-based marginal regression model

In this section, we introduce a copula-based regression model with exponentially distributed marginals based on the Fréchet copula, and investigate D-optimal designs with global and local approaches for this model.

4.1 The model

Probability models with positive support have been used a lot in scientific studies due to their good fit with real-world random events, especially when the variable under study is defined in terms of time. For example, in areas such as reliability, survival analysis, and growth regression models, etc. One of the most popular and widely used models is the exponential distribution. The exponential distribution is a continuous distribution supported on $[0, \infty)$ with a probability density and cumulative distribution functions defined as follows

$$\begin{aligned} f_Y(y) &= \lambda e^{-\lambda y}, \quad y > 0, \\ F_Y(y) &= 1 - e^{-\lambda y}, \end{aligned}$$

where $\lambda > 0$ is the (rate) parameter of the exponential distribution and $E[Y] = \frac{1}{\lambda}$. For more details regarding the exponential distribution and its statistical properties, we refer to Balakrishnan and Nevzorov (2003).

Regression models are typically derived to model the parameters (which describe the mean, median, mode, etc.) of a probability distribution. Here, this approach is applied to the exponential distribution. Let $\boldsymbol{x} = (x_1, \ldots, x_k)^T$ be the vector of covariates. The parameter λ is linked to the covariates by the logarithmic link function $\log \lambda = \boldsymbol{x}^T \boldsymbol{\theta}$, where $\boldsymbol{\theta} = (\theta_1, \ldots, \theta_k)^T$ is the vector of unknown regression coefficients.

Therefore, the mean of Y is $E[Y|\boldsymbol{x}] = E[Y(\boldsymbol{x})] = \eta(\boldsymbol{x}, \boldsymbol{\theta}) = \frac{1}{\exp(\boldsymbol{x}^T \boldsymbol{\theta})}$ and the predicted density function and distribution function of Y for a given $\boldsymbol{x}$ are defined as follows, respectively

$$f(y|\boldsymbol{x}) = f_Y(y(\boldsymbol{x}, \boldsymbol{\theta})) = e^{\boldsymbol{x}^T \boldsymbol{\theta}} \exp\left(-e^{\boldsymbol{x}^T \boldsymbol{\theta}} y\right), \tag{6}$$

$$F(y|\boldsymbol{x}) = F_Y(y(\boldsymbol{x}, \boldsymbol{\theta})) = 1 - \exp\left(-e^{\boldsymbol{x}^T \boldsymbol{\theta}} y\right). \tag{7}$$

Now, let Y_1 and Y_2 be both exponentially distributed, and suppose that we are interested in considering the influence of a covariate vector $\boldsymbol{x}$ on the dependence structure of a vector of interest $(Y_1, Y_2)^T$. Obviously, a copula function is used to illustrate how the relationship between Y_1 and Y_2 varies with the influence of a covariate vector $\boldsymbol{x}$. To describe this copula function, considering the joint distribution function of $(Y_1, Y_2)^T$ given covariate vector $\boldsymbol{x}$ as $F_{Y_1,Y_2}(y_1(\boldsymbol{x}, \boldsymbol{\theta}), y_2(\boldsymbol{x}, \boldsymbol{\theta}))$, if the marginal distribution functions of Y_1 and Y_2 given covariate vector $\boldsymbol{x}$ are denoted by $F_{Y_1}(y_1(\boldsymbol{x}, \boldsymbol{\theta}))$ and $F_{Y_2}(y_2(\boldsymbol{x}, \boldsymbol{\theta}))$, respectively. Then Sklar's theorem ensures that there exists a unique copula $C : [0, 1] \times [0, 1] \to [0, 1]$, such that (use also (7))

$$F_{Y_1,Y_2}(y_1(\boldsymbol{x}, \boldsymbol{\theta}), y_2(\boldsymbol{x}, \boldsymbol{\theta}); \boldsymbol{\alpha}) = C\left(1 - \exp\left(-e^{\boldsymbol{x}^T \boldsymbol{\theta}} y_1\right), 1 - \exp\left(-e^{\boldsymbol{x}^T \boldsymbol{\theta}} y_2\right); \boldsymbol{\alpha}\right),$$

where $\boldsymbol{\alpha} = (\alpha_1, \alpha_2, \ldots, \alpha_z)^T$ is the vector of unknown copula parameters if the copula function is of parametric type.

Thus, then the corresponding joint density function $f_{Y_1,Y_2}\left(y_1(\boldsymbol{x},\boldsymbol{\theta}), y_2(\boldsymbol{x},\boldsymbol{\theta}); \boldsymbol{\alpha}\right)$ is obtained as (use also (6))

$$\begin{aligned} f_{Y_1,Y_2}\left(y_1(\boldsymbol{x},\boldsymbol{\theta}), y_2(\boldsymbol{x},\boldsymbol{\theta}); \boldsymbol{\alpha}\right) = & \, c\left(1-\exp\left(-e^{\boldsymbol{x}^T\boldsymbol{\theta}}y_1\right), 1-\exp\left(-e^{\boldsymbol{x}^T\boldsymbol{\theta}}y_2\right); \boldsymbol{\alpha}\right) \\ & \times \left(e^{\boldsymbol{x}^T\boldsymbol{\theta}}\exp\left(-e^{\boldsymbol{x}^T\boldsymbol{\theta}}y_1\right)\right)\left(e^{\boldsymbol{x}^T\boldsymbol{\theta}}\exp\left(-e^{\boldsymbol{x}^T\boldsymbol{\theta}}y_2\right)\right). \quad (8) \end{aligned}$$

We will discuss the optimal designs for the above copula-based model by assuming that the dependent structure of the random pair (Y_1, Y_2) is influenced by only one covariate x. Indeed, knowing how the dependence structure changes with the value taken by one covariate is an interesting topic in the relevant literature. An example is given in Gijbels et al. (2011), where a copula function is used to illustrate how the relationship between the life expectancy of men (Y_1) and women (Y_2) varies with the growth of domestic product $X = x$. Therefore, for each design point x from the design space $\boldsymbol{\chi}$, we may observe a pair of exponential random variables Y_1 and Y_2, such that $E[Y_i(x)] = \frac{1}{\exp(\theta_i x)}$, $i = 1, 2$. Hence, (8) reduces to the following:

$$\begin{aligned} f_{Y_1,Y_2}\left(y_1(x,\theta_1), y_2(x,\theta_2); \boldsymbol{\alpha}\right) \;=\; & c\left(1-\exp\left(-e^{\theta_1 x}y_1\right), 1-\exp\left(-e^{\theta_2 x}y_2\right); \boldsymbol{\alpha}\right) \\ & \times \left(e^{\theta_1 x}\exp\left(-e^{\theta_1 x}y_1\right)\right)\left(e^{\theta_2 x}\exp\left(-e^{\theta_2 x}y_2\right)\right). \quad (9) \end{aligned}$$

In the upcoming subsection, optimal designs are obtained for the copula-based model in (9).

Remark 4.1. *Note that since the number of covariates is one, it is expected that in the optimal designs that will be obtained, the design points will all have the same value with equal weights (i.e., balance designs with fixed optimal points). This will be confirmed in the numerical results that will be carried out in the next subsection.*

4.2 Design of experiments

D-optimal design. In the following, the D-optimal design has been obtained according to the model (9) based on the Fréchet copula. Since the Fréchet copula in (5) is a parametric copula, so the optimal design is focused on the marginals parameters and the copula parameters. Using the Fréchet copula (5) in model (9), we have the following joint density function:

$$f_{Y_1,Y_2}\left(y_1(x,\theta_1), y_2(x,\theta_2); \gamma\right) = \gamma\left(e^{\theta_1 x}\exp\left(-e^{\theta_1 x}y_1\right)\right)\left(e^{\theta_2 x}\exp\left(-e^{\theta_2 x}y_2\right)\right). \quad (10)$$

Using (1), the Fisher information matrix for this copula-based model is obtained as follows:

$$M(x,\theta_1,\theta_2,\gamma) = diag\left(x^2, x^2, \frac{1}{\gamma^2}\right).$$

Since there are three parameters in the non-linear copula-based model (10), so the number of points in the optimal design satisfies in the inequality $3 \le r \le 6$ (see Subsection 2.2). Thus, by using (3), the information matrix of the design $\xi = \left\{ \begin{smallmatrix} x_1 & x_2 & \ldots & x_r \\ w_1 & w_2 & \ldots & w_r \end{smallmatrix} \right\}$,

$3 \leq r \leq 6$, is $\mathcal{M}(\xi, \theta_1, \theta_2, \gamma) = \sum_{i=1}^{r} w_i M(x_i, \theta_1, \theta_2, \gamma)$, where $\sum_{i=1}^{r} w_i = 1$. Therefore, the D-optimal criterion, that is, $\psi(\mathcal{M}(\xi, \theta_1, \theta_2, \gamma)) = \log(\det(\mathcal{M}(\xi, \theta_1, \theta_2, \gamma)))$, is obtained as follows:

$$\psi(\mathcal{M}(\xi, \theta_1, \theta_2, \gamma)) = \log\left(\frac{\left(\sum_{i=1}^{r} w_i x_i^2\right)^2}{\gamma^2}\right) = \log\left(\sum_{i=1}^{r} w_i x_i^2\right)^2 - \log(\gamma^2). \tag{11}$$

From this, it is clear that the value of γ does not affect the optimization process. So, if we consider the local approach or the Bayesian approach, the D-optimal criterion in (11) reduces to the formula $\log\left(\sum_{i=1}^{r} w_i x_i^2\right)^2 + L$, where L is a real number. Therefore, the optimal designs that will be obtained are global with respect to the design space that is considered. The obtained optimal designs are summarized in Table 1.

Table 1: D-optimal designs for the copula-based model (10).

Three-point				
		ξ^*		
Design space	Design	First	Second	Thired
$\chi = [a, b]$	x^*	$\frac{a+b}{2}$	$\frac{a+b}{2}$	$\frac{a+b}{2}$
$\forall a, b > 0$	w^*	0.333	0.333	0.333
$\chi = [a, \infty)$	x^*	1	1	1
$0 \leq a \leq 1$	w^*	0.333	0.333	0.333
$\chi = [a, \infty)$	x^*	a	a	a
$\forall a > 1$	w^*	0.333	0.333	0.333

Four-point					
		ξ^*			
Design space	Design	First	Second	Thired	Fourth
$\chi = [a, b]$	x^*	$\frac{a+b}{2}$	$\frac{a+b}{2}$	$\frac{a+b}{2}$	$\frac{a+b}{2}$
$\forall a, b > 0$	w^*	0.250	0.250	0.250	0.250
$\chi = [a, \infty)$	x^*	1	1	1	1
$0 \leq a \leq 1$	w^*	0.250	0.250	0.250	0.250
$\chi = [a, \infty)$	x^*	a	a	a	a
$\forall a > 1$	w^*	0.250	0.250	0.250	0.250

Five-point						
		ξ^*				
Design space	Design	First	Second	Thired	Fourth	Fifth
$\chi = [a, b]$	x^*	$\frac{a+b}{2}$	$\frac{a+b}{2}$	$\frac{a+b}{2}$	$\frac{a+b}{2}$	$\frac{a+b}{2}$
$\forall a, b > 0$	w^*	0.200	0.200	0.200	0.200	0.200
$\chi = [a, \infty)$	x^*	1	1	1	1	1
$0 \leq a \leq 1$	w^*	0.200	0.200	0.200	0.200	0.200
$\chi = [a, \infty)$	x^*	a	a	a	a	a
$\forall a > 1$	w^*	0.200	0.200	0.200	0.200	0.200

Six-point							
		ξ^*					
Design space	Design	First	Second	Thired	Fourth	Fifth	Sixth
$\chi = [a, b]$	x^*	$\frac{a+b}{2}$	$\frac{a+b}{2}$	$\frac{a+b}{2}$	$\frac{a+b}{2}$	$\frac{a+b}{2}$	$\frac{a+b}{2}$
$\forall a, b > 0$	w^*	0.166	0.166	0.166	0.166	0.166	0.166
$\chi = [a, \infty)$	x^*	1	1	1	1	1	1
$0 \leq a \leq 1$	w^*	0.166	0.166	0.166	0.166	0.166	0.166
$\chi = [a, \infty)$	x^*	a	a	a	a	a	a
$\forall a > 1$	w^*	0.166	0.166	0.166	0.166	0.166	0.166

Remark 4.2. *It is observed from Table 1 that the optimal points obtained in each case are all the same with equal weights. The reason was already explained in Remark 4.1. This is a special situation, and in the relevant literature, one of the points with a weight of 1 is usually written to represent such an optimal design. For example, if we consider* $\boldsymbol{\chi} = [2, 6]$*, then by Table 1, the three-point optimal design is* $\xi^* = \left\{ {4 \atop 0.333}\ {4 \atop 0.333}\ {4 \atop 0.333} \right\}$. *Thus, it is written as* $\xi^* = \left\{ {4 \atop 1} \right\}$.

D**-efficiency.** The D-efficiency of an r-point design $\xi = \left\{ {x \atop w}\ {x \atop w}\ {\cdots \atop \cdots}\ {x \atop w} \right\} = \left\{ {x \atop 1} \right\}$ with respect to the r-point D-optimal design $\xi^* = \left\{ {x^* \atop w^*}\ {x^* \atop w^*}\ {\cdots \atop \cdots}\ {x^* \atop w^*} \right\} = \left\{ {x^* \atop 1} \right\}$ is given by

$$D_{eff}(\xi^*, \xi) = \left(\frac{\det(\mathcal{M}(\xi, \theta_1, \theta_2, \gamma))}{\det(\mathcal{M}(\xi^*, \theta_1, \theta_2, \gamma))} \right)^{\frac{1}{3}} = \left(\frac{\frac{\left(\sum_{i=1}^{r} w_i x_i^2\right)^2}{\gamma^2}}{\frac{\left(\sum_{i=1}^{r} w_i{}^* x_i^*{}^2\right)^2}{\gamma^2}} \right)^{\frac{1}{3}} = \left(\frac{x}{x^*} \right)^{\frac{4}{3}}. \tag{12}$$

Therefore, by using Table 1, for $3 \le r \le 6$, we have

- If $\boldsymbol{\chi} = [a, b]$ and $a, b > 0$, then $D_{eff}(\xi^*, \xi) = \left(\frac{2x}{a+b} \right)^{\frac{4}{3}}$;
- If $\boldsymbol{\chi} = [a, \infty]$ and $0 \le a \le 1$, then $D_{eff}(\xi^*, \xi) = x^{\frac{4}{3}}$;
- If $\boldsymbol{\chi} = [a, \infty]$ and $a > 1$, then $D_{eff}(\xi^*, \xi) = \left(\frac{x}{a} \right)^{\frac{4}{3}}$.

As an example, let us consider $\boldsymbol{\chi} = [1, 5]$ and $\xi = \left\{ {2 \atop 1} \right\}$. In this case, the D-efficiency of the design ξ with respect to ξ^* is $D_{eff}(\xi^*, \xi) = 0.8735$. The plot of the D-efficiency in (12) for design space $\boldsymbol{\chi} = [a, b]$ is depicted in Figure 2. It can be seen from Figure 2 that increasing a leads to a decrease in the D-efficiency.

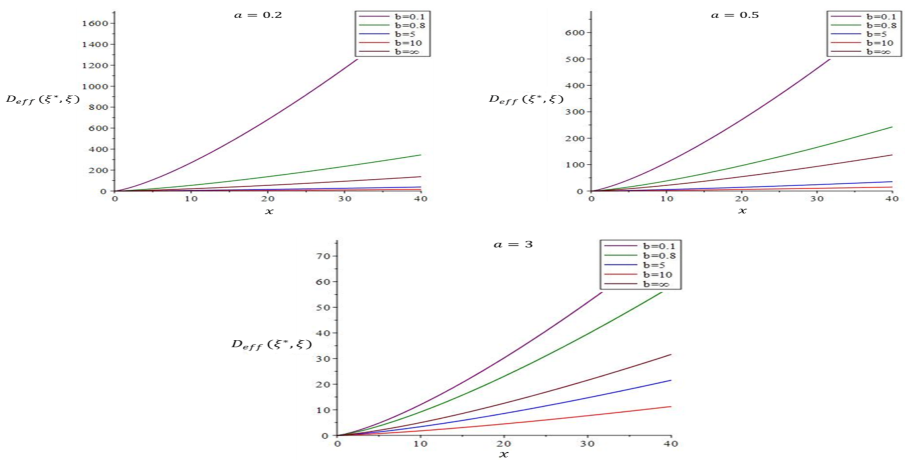

Figure 2: Plots of $D_{eff}(\xi^*, \xi)$ in (12) versus x for the design space $\boldsymbol{\chi} = [a, b]$ for $a = 0.2, 0.3, 3$, and some different values of b.

5 Conclusions

In this paper, the optimal experimental design for a copula marginal regression model with exponential marginals is discussed. More precisely, we considered the dependence structure of a random pair from an exponential distribution conditionally upon only one covariate as a regression model, then investigated the D-optimal design for this regression model. Since the dependency structure in the model was expressed by a copula, the considered model was a copula-based model. The copula we used was the Fréchet copula. The numerical results showed that the optimal designs all had general forms according to the considered design space. Optimal design for copula-based regression models has been introduced and developed in recent years, and therefore this paper can be useful for further developments.

Acknowledgements

The authors would like to thank the Editor and three anonymous reviewers for their valuable suggestions and comments that helped improve the original version of the paper.

References

Abdollahi N.A., Jafari, H. and Khazaei, S. (2024). Locally, Bayesian and non parametric Bayesian optimal designs for unit exponential regression model, *Communications in Statistics-Theory and Methods*. https://doi.org/10.1080/03610926.2024.2328182

Aminnejad, M. and Jafari, H. (2017). Bayesian A- and D-optimal designs for gamma regression model with inverse link function. *Communications in Statistics-Simulation and Computation*, **46**(10): 8166–8189.

Atkinson, A.C., Donev, A.N. and Tobias, R.D. (2007). *Optimum Experimental Design, with SAS*. Oxford: Oxford University Press.

Balakrishnan, N. and Nevzorov, V.B. (2003). *A Primer on Statistical Distributions*. New Jersey: John Wiley & Sons.

Burghaus, I. and Dette, H. (2014). Optimal designs for nonlinear regression models with respect to non-informative priors. *Journal of Statistical Planning and Inference*, **154**:12–25.

Burkner, P.C., Schwabe, R. and Holling, H. (2019). Optimal designs for the generalized partial credit model. *British Journal of Mathematical and Statistical Psychology*, **72**(2):271–293.

Chaloner, K.M. and Duncan, G.T. (1983). Assessment of a beta prior distribution: PM elicitation. *Journal of the Royal Statistical Society. Series D (The Statistician)*, **32**(1-2):174–180.

Cherubini, U., Gobbi, F. and Mulinacci, S. (2016). *Convolution Copula Econometrics*. Cham: Springer.

Cherubini, U., Luciano, E. and Vecchiato, W. (2004). *Copula Methods in Finance*. John Wiley & Sons.

Chernoff, H. (1953). Locally optimal designs for estimating parameters. *The Annals of Mathematical Statistics*, **24**(4):586–602.

Deldossi, L., Osmetti, S.A. and Tommasi, C. (2018). Optimal design to discriminate between rival copula models for a bivariate binary response. *TEST*, **28**:147–165.

Denman, N.G., McGree, J.M., Eccleston, J.A. and Duffull, S.B. (2011). Design of experiments for bivariate binary responses modelled by Copula functions. *Computational Statistics and Data Analysis*, **55**(4):1509–1520.

Dette, H., Melas, V.B. and Wong, W.K. (2006). Locally D-optimal designs for exponential regression. *Statistica Sinica*, **16**(3):789–803.

Dette, H. and Neugebauer, H.-M. (1997). Bayesian D-optimal designs for exponential regression models. *Journal of Statistical Planning and Inference*, **60**(2):331–349.

Durante, F. and Perrone, E. (2016). Asymmetric copulas and their application in design of experiments. In: Saminger-Platz S., Mesiar, R. (eds.) *On Logical, Algebraic, and Probabilistic Aspects of Fuzzy Set Theory.* Vol. 336, 157–172, Cham: Springer.

Durante, F. and Sempi, C. (2015). *Principles of Copula Theory.* Vol. 474, Boca Raton, FL: CRC press.

Fedorov, V. and Hackl, P. (2012). *Model-oriented Design of Experiments.* New York: Springer.

Firth, D. and Hinde, J.P. (1997). On Bayesian D-optimum design criteria and the equivalence theorem in non-linear models. *Journal of the Royal Statistical Society Series B: Statistical Methodology*, **59**(4), 793–797.

Gijbels, I., Veraverbeke, N. and Omelka, M. (2011). Conditional copulas, association measures and their applications. *Computational Statistics & Data Analysis*, **55**(5):1919–1932.

Goudarzi, M., Jafari, H. and Khazaei, S. (2019). Nonparametric Bayesian optimal designs for exponential regression model. *Communications in Statistics-Simulation and Computation*, **50**(6):1809–1819.

Graßhoff, U., Holling, H. and Schwabe, R. (2012). Optimal designs for the Rasch model. *Psychometrika*, **77**(4):710–723.

Joag-Dev, K. and Proschan, P. (1983). Negative Association of Random Variables with Applications. *The Annals of Statistics*, **11**(1):286–295.

Joe, H. (1990). Multivariate concordance. *Journal of Multivariate Analysis*, **35**(1):12–30.

Kiefer, J. (1959). Optimum experimental design. *Journal of the Royal Statistical Society Series B: Statistical Methodology*, **21**(2):272–319.

Kiefer, J. and Wolfowitz, J. (1959). Optimum designs in regression problems. *The Annals of Mathematical Statistics*, **30**(2):271–294.

Kiefer, J. and Wolfowitz, J. (1960). The equivalence of two extremum problems. *Canadian Journal of Mathematics*, **12**:363–366.

Kiefer, J. (1974). General equivalence theory for optimum designs (approximate theory). *The Annals of Statistics*, **2**(5):849–79.

Lehmann, E.L. (1966). Some Concepts of Dependence. *The Annals of Statistics*, **37**(5):1137–1153.

Li, J., Bárdossy, A., Guenni, L., and Liu, M. (2011). A copula based observation network design approach. *Environmental Modelling & Software*, **26**(11):1349–1357.

Mukhopadhyay, S. and Haines, L.M. (1995). Bayesian D-optimal designs for the exponential growth model. *Journal of Statistical Planning and Inference*, **44**(3):385–397.

Nelsen, R.B. (2006). *An Introduction to Copulas.* New York: Springer.

Parsa Maram, P. and Jafari, H. (2016). Bayesian D-optimal design for logistic regression model with exponential distribution for random intercept. *Journal of Statistical Computation and Simulation*, **86**(10):1856–1868.

Perrone, E. and Müller, W.G. (2016). Optimal designs for Copula models. *Statistics*, **50**(4):917–929.

Perrone, E., Rappold, A. and Müller, W.G. (2016). D_s-optimality in copula models. *Statistical Methods & Applications*, **26**(3):403–418.

Pilz, J., Kazianka, H. and Spöck, G. (2012). Some advances in Bayesian spatial prediction and sampling design. *Spatial Statistics*, **1**:65–81.

Pukelsheim, F. (1993). *Optimal Design of Experiments.* New York: John Wiley.

Rappold, A., Müller, W.G. and Woods, D.C. (2020). Copula-based robust optimal block designs. *Applied Stochastic Models in Business and Industry*, **36**(1):210–219.

Salvadori, G., De Michele, C., Kottegoda, N.T. and Rosso, R. (2007). *Extremes in Nature: An Approach Using Copulas.* Dordrecht: Springer.

Schmidt, R., Faldum, A., Witt, O. and Gerss, J. (2014). Adaptive designs with arbitrary dependence structure. *Biom J*, **56**(1):86–106.

Silvey, S.D. (1980). *Optimal Design.* London: Chapman and Hall.

Sklar, A. (1959). Fonctions de répartition á n dimensions et leurs marges, *Publications de l'institue de Statistique de l'Universitte de Paris*, **8**(3):229–231.

Smith, K. (1918). On the standard deviations of adjusted and interpolated values of an observed polynomial function and it's constant and the guidance they give towards a proper choice of the distribution of observations. *Biometrika*, **12**(1/2):1–85.

Xie, Z., Wang, F., Yang, Y. and Guo, N. (2022). Multivariate Bernstein Fréchet copulas. *Statistics and Its Interface.* **15**(4):527–547.