Article

Copula-Based Estimation Methods for a Common Mean Vector for Bivariate Meta-Analyses

Jia-Han Shih [1], Yoshihiko Konno [2], Yuan-Tsung Chang [3] and Takeshi Emura [4],*

[1] Institute of Statistical Science, Academia Sinica, Taipei 11529, Taiwan; jhshih@stat.sinica.edu.tw
[2] Department of Mathematical and Physical Sciences, Japan Women's University, Tokyo 112-8681, Japan; konno@fc.jwu.ac.jp
[3] Department of Social Information, Mejiro University, Tokyo 161-8539, Japan; chogenso@gmail.com
[4] Biostatistics Center, Kurume University, Kurume, Fukuoka 830-0011, Japan
* Correspondence: takeshiemura@gmail.com

Abstract: Traditional bivariate meta-analyses adopt the bivariate normal model. As the bivariate normal distribution produces symmetric dependence, it is not flexible enough to describe the true dependence structure of real meta-analyses. As an alternative to the bivariate normal model, recent papers have adopted "copula" models for bivariate meta-analyses. Copulas consist of both symmetric copulas (e.g., the normal copula) and asymmetric copulas (e.g., the Clayton copula). While copula models are promising, there are only a few studies on copula-based bivariate meta-analysis. Therefore, the goal of this article is to fully develop the methodologies and theories of the copula-based bivariate meta-analysis, specifically for estimating the common mean vector. This work is regarded as a generalization of our previous methodological/theoretical studies under the FGM copula to a broad class of copulas. In addition, we develop a new R package, "*CommonMean.Copula*", to implement the proposed methods. Simulations are performed to check the proposed methods. Two real dataset are analyzed for illustration, demonstrating the insufficiency of the bivariate normal model.

1. Introduction

Bivariate outcomes often arise in meta-analyses on scientific studies, such as education and medicine. Educational researchers may analyze bivariate exam scores on verbal and mathematics [1,2], or on mathematics and statistics [3]. Medical experts may analyze bivariate risk scores on myocardial infection and cardiovascular death for diabetes patients [4,5]. Bivariate meta-analyses are statistical methods designed for these meta-analytical studies [6]. Dependence between two outcomes should be considered while performing bivariate meta-analyses. If one simply considers univariate (marginal) analysis for each outcome separately, any possible dependence between the outcomes is ignored. Riley [2] and Copas et al. [7] showed that ignoring the dependence between two outcomes increases the error for estimating parameters due to the loss of information. In medical research, dependence itself can be of clinical importance, e.g., dependence between two survival outcomes in meta-analysis [8–11].

In the traditional bivariate meta-analyses, the parameters of interest are the means of a bivariate normal model [6]. However, the bivariate normal model is not flexible enough to describe the true dependence structure of real meta-analyses. It will be shown that the bivariate normal mode fits poorly to the dependence structure of real bivariate meta-analyses (Section 8). This has motivated researchers to consider alternative models.

As an alternative to the bivariate normal model, recent papers have adopted "copula" models for bivariate meta-analyses [3,5,12–15]. Copula models are flexible as they allow a variety of dependence structures. Copulas consist of both symmetric copulas (e.g., the normal

copula) and asymmetric copulas (e.g., the Clayton copula). Copula models have become very popular in all areas of science by replacing the traditional multivariate normal models. In astronomy, Takeuchi [16] constructed the bivariate luminosity density functions using the FGM copula; see reference [17] for the application of the FGM copula to engineering. In ecology, Ghosh et al. [18] applied copulas to model the dependence structure in environmental and biological variables. In environmental science, Alidoost et al. [19] used bivariate copulas in the analysis of temperature. See the survey of [20] for applications to energy, forestry, and environmental sciences. The books of [21,22] are devoted to the applications of copulas in survival analysis; see also references [11,23–25].

While bivariate copula models for meta-analyses are promising, there are only a few methodologically and theoretically solid studies on copula-based bivariate meta-analysis. For instance, the detailed theoretical studies of [3] are limited to the FGM copula. Other copula-based meta-analyses published in biostatistical journals, such as [5,12–15], are proposed without theoretical details. Furthermore, copula-based bivariate meta-analyses have not been implemented in a free software environment.

Therefore, the goal of this article is to fully develop the methodologies and theories of the copula-based bivariate meta-analysis for estimating the common mean vector. This work is regarded as a large generalization of our previous methodological/theoretical studies under the FGM copula model [3] to a broad class of copula models. In this article, we obtain theoretical results, including the formula of the information matrix and large sample theories. Our theoretical results guarantee the applications of many copulas, such as the Clayton, Gumbel, Frank, and normal copulas, in addition to the FGM copula. In addition, we developed a new R package, "*CommonMean.Copula*" [26], to implement the proposed methods under the five copulas. Therefore, the aim of the article is to make a solid development of the methodologies, theories, and practical implementations of copula-based bivariate meta-analysis for the common mean, which are not yet available in the literature.

The article is organized as follows. Section 2 reviews the background of this research. Section 3 introduces the proposed model and estimator. Section 4 provides the asymptotic theory and Section 5 gives confidence sets. Section 6 introduces our new R package. Section 7 conducts simulations to check the accuracy of the proposed methods. Section 8 analyzes two real datasets for illustration. Section 9 extends the proposed methods to non-normal data. Finally, Section 10 concludes with a discussion.

2. Background

This section reviews the literature on bivariate meta-analyses and the concept of copulas.

2.1. Bivariate Meta-Analysis

We review the bivariate meta-analysis method for bivariate continuous outcomes [6,27]. For each study i, let the bivariate outcomes, Y_{i1} and Y_{i2}, follow a bivariate normal distribution

$$\mathbf{Y}_i = \begin{bmatrix} Y_{i1} \\ Y_{i2} \end{bmatrix} \sim N\left(\boldsymbol{\mu} = \begin{bmatrix} \mu_1 \\ \mu_2 \end{bmatrix}, \boldsymbol{\Omega}_i = \begin{bmatrix} \sigma_{i1}^2 & \rho_i \sigma_{i1} \sigma_{i2} \\ \rho_i \sigma_{i1} \sigma_{i2} & \sigma_{i2}^2 \end{bmatrix}\right), \quad i = 1, 2, \ldots, n, \quad (1)$$

where $\rho_i \in (-1, 1)$ is the within-study correlation for each i. In Equation (1), all the responses ($\mathbf{Y}_i$s) share the common mean vector ($\boldsymbol{\mu}$). The covariance matrix $\boldsymbol{\Omega}_i$ is assumed to be known (from the i-th study) in usual bivariate meta-analyses. We do not consider a setting where the covariance is unknown [28,29].

Then, the MLE of the common mean vector is quite easily computed as

$$\hat{\boldsymbol{\mu}}_n^{\text{Normal}} = \begin{bmatrix} \hat{\mu}_{n,1}^{\text{Normal}} \\ \hat{\mu}_{n,2}^{\text{Normal}} \end{bmatrix} = \left(\sum_{i=1}^{n} \boldsymbol{\Omega}_i^{-1}\right)^{-1} \sum_{i=1}^{n} \boldsymbol{\Omega}_i^{-1} \mathbf{Y}_i.$$

One could use the R package *mvmeta* [30], although the above computation is easy.

The bivariate normal model (1) does not allow for a different dependence structure between the two outcomes. In practice, the bivariate normal model (1) can be too restrictive, as there are various dependence patterns between two outcomes. For example, to model the luminosity function of galaxies, Takeuchi [16] pointed out that the FGM copula model offers a more ideal shape than the normal copula model from a physical point of view. Such a limitation motivates us to construct a *general* copula model that can describe various dependence structures.

2.2. Copulas

This subsection prepares the basic terms on copulas that will subsequently be used.

A copula is a bivariate distribution function whose margins are uniformly distributed on the unit interval [31,32]. Copulas are indispensable tools when modelling a dependence structure between two random variables. We specifically consider the following *parametric* copulas.

The normal copula: The copula function is

$$C_\rho^{\text{Normal}}(u,v) = \Phi_\rho\left\{\Phi^{-1}(u), \Phi^{-1}(v)\right\}, \quad -1 < \rho < 1, \quad 0 < u, v < 1,$$

where $\Phi_\rho(\cdot, \cdot)$ is the cumulative distribution function (CDF) of the bivariate standard normal distribution with correlation ρ and Φ^{-1} is the inverse of the standard normal CDF Φ. While this copula is easy to understand, it has a complex form involving two implicit functions Φ_ρ and Φ^{-1}. The following two copulas provide simpler forms than the normal copula.

The Farlie–Gumbel–Morgenstern (FGM) copula [33]**:** The copula function is

$$C_\theta^{\text{FGM}}(u,v) = uv\{1 + \theta(1-u)(1-v)\}, \quad -1 \le \theta \le 1, 0 < u, v < 1.$$

The FGM copula has a very simple form, and is a fundamental copula, which has been extended to a variety of copulas, called the generalized FGM copulas [34–38].

The Clayton copula [39]**:** The copula function is

$$C_\alpha^{\text{Clayton}}(u,v) = (u^{-\alpha} + v^{-\alpha} - 1)^{-1/\alpha}, \quad \alpha > 0, 0 < u, v < 1.$$

The Clayton copula is one of the simplest and most frequently used copulas in applications. The Clayton copula is derived from the gamma frailty model, leading to its remarkable popularity in survival data analysis [22,40]. It has a lower tail dependence [31], but is not tractable for modeling negative dependence.

The Gumbel copula [41]**:** The copula function is

$$C_\beta^{\text{Gumbel}}(u,v) = \exp\left[-\left\{(-\log u)^\beta + (-\log v)^\beta\right\}^{1/\beta}\right], \quad \beta \ge 1, 0 < u, v < 1.$$

The Gumbel copula is a popular copula with upper tail dependence [31]. The Gumbel copula does not offer a negative dependence, as in the Clayton copula.

The Frank copula [42]**:** The copula function is

$$C_\gamma^{\text{Frank}}(u,v) = -\frac{1}{\gamma}\log\left\{1 + \frac{(e^{-\gamma u} - 1)(e^{-\gamma v} - 1)}{e^{-\gamma} - 1}\right\}, \quad \gamma \ne 0, 0 < u, v < 1.$$

The Frank copula does not have tail dependence [31]. Unlike the Clayton and Gumbel copulas, it can model both positive and negative dependences as the normal copula.

Under the null parameter (e.g., $\theta = 0$), all the above copulas reduce to the independence copula $\Pi(u,v) = uv$. As the parameter departs from the null, the dependence gets stronger.

We define the notations for partial derivatives (if they exist) as

$$C^{[j,k]}(u,v) = \frac{\partial^{j+k}}{\partial u^j \partial v^k} C(u,v); \quad j,k \in \{0,1,2,\ldots\}.$$

For instance,

$$C^{[1,0]}(u,v) = \frac{\partial}{\partial u} C(u,v), \quad C^{[0,1]}(u,v) = \frac{\partial}{\partial v} C(u,v), \quad C^{[1,1]}(u,v) = \frac{\partial^2}{\partial u \partial v} C(u,v),$$

where $C^{[1,1]}$ is called the copula density.

The copula is symmetric if $C^{[1,1]}(u,v) = C^{[1,1]}(1-u, 1-v)$. This means that the normal and FGM copulas are symmetric while the Clayton and Gumbel copulas are asymmetric. This symmetry should not be confused with the exchangeability $C(u,v) = C(v,u)$. All the aforementioned parametric copulas are exchangeable.

3. Proposed Methods

This section proposes a general copula-based approach for estimating a bivariate common mean vector. We first define the bivariate copula model and provide sufficient conditions for the copula parameter to be identifiable. We then develop a maximum likelihood estimator (MLE) for the common mean vector. In addition, we derive the expression for the information matrix.

3.1. General Copula Model for the Common Mean

This subsection proposes a new model for estimating the common mean in bivariate meta-analyses.

For $i = 1, 2, \ldots, n$, let $\mathbf{Y}_i = (Y_{i1}, Y_{i2})$ be a random vector satisfying

$$Y_{i1} \sim N\left(\mu_1, \sigma_{i1}^2\right), Y_{i2} \sim N\left(\mu_2, \sigma_{i2}^2\right), \boldsymbol{\mu} \equiv E(\mathbf{Y}_i) = \begin{bmatrix} E(Y_{i1}) \\ E(Y_{i2}) \end{bmatrix} = \begin{bmatrix} \mu_1 \\ \mu_2 \end{bmatrix},$$

$$\boldsymbol{\Omega}_i \equiv Cov(\mathbf{Y}_i) = \begin{bmatrix} Var(Y_{i1}) & Cov(Y_{i1}, Y_{i2}) \\ Cov(Y_{i1}, Y_{i2}) & Var(Y_{i2}) \end{bmatrix} = \begin{bmatrix} \sigma_{i1}^2 & \rho_i \sigma_{i1} \sigma_{i2} \\ \rho_i \sigma_{i1} \sigma_{i2} & \sigma_{i2}^2 \end{bmatrix}.$$

Here, we call $\boldsymbol{\mu} = (\mu_1, \mu_2)$ the 'common mean vector' since it is common across $i = 1, 2, \ldots, n$. Our target is the estimation of $\boldsymbol{\mu}$ when $\boldsymbol{\Omega}_i$, $i = 1, 2, \ldots, n$ are known. In general, $\boldsymbol{\Omega}_i \neq \boldsymbol{\Omega}_j$ for some $i \neq j$, and, therefore, the random vectors $\mathbf{Y}_i$, $i = 1, 2, \ldots, n$ are independent but not identically distributed (i.n.i.d.). While the marginal normality is specified, the bivariate normality is unspecified. We only specify the equation $\text{Corr}(Y_{i1}, Y_{i2}) = \rho_i$, where ρ_i is known.

We now specify a bivariate distribution for $\mathbf{Y}_i$. According to Sklar's Theorem [43], for copulas C_{θ_i}, $i = 1, 2, \ldots, n$, we define the bivariate CDFs

$$Pr(Y_{i1} \leq y_1, Y_{i2} \leq y_2) = C_{\theta_i}\left\{\Phi\left(\frac{y_1 - \mu_1}{\sigma_{i1}}\right), \Phi\left(\frac{y_2 - \mu_2}{\sigma_{i2}}\right)\right\}, i = 1, 2, \ldots, n.$$

However, since ρ_i is known, the copula can be restricted. To see the problem clearly, we define the *correlation function* $\rho_C : \Theta \mapsto R_C$ as

$$\rho_C(\boldsymbol{\theta}) = E\left\{\left(\frac{Y_{i1} - \mu_1}{\sigma_{i1}}\right)\left(\frac{Y_{i2} - \mu_2}{\sigma_{i2}}\right)\right\} = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} z_1 z_2 dC_{\boldsymbol{\theta}}\{\Phi(z_1), \Phi(z_2)\},$$

where $R_C \equiv \{\rho_C(\boldsymbol{\theta}) : \boldsymbol{\theta} \in \Theta\}$ denotes the range of ρ_C that depends on the choice of $C_{\boldsymbol{\theta}}$. The correlation function ρ_C does not depend on $\boldsymbol{\mu}$. For the copula to be useful in real meta-analyses, $\boldsymbol{\theta}_i$ has to be identifiable from ρ_i. This means that one has to be able to solve the equation $\rho_C(\boldsymbol{\theta}) = \rho$. Now, we define our general copula model for a bivariate common mean vector.

Definition 1. *(Copula-based common mean model): The copula-based common mean model is*

$$Pr(Y_{i1} \leq y_1, Y_{i2} \leq y_2) = C_{\boldsymbol{\theta}_i}\left\{\Phi\left(\frac{y_1 - \mu_1}{\sigma_{i1}}\right), \Phi\left(\frac{y_2 - \mu_2}{\sigma_{i2}}\right)\right\}, \quad i = 1, 2, \ldots, n, \qquad (2)$$

where the copula parameter $\boldsymbol{\theta}_i$ is identified by $\rho_C(\boldsymbol{\theta}_i) = \rho_i$ for $i = 1, 2, \ldots, n$.

To explain the flexibility and generality of our model, we give examples for $C_{\boldsymbol{\theta}_i}$.

Example 1. *(the normal copula): Under the normal copula, the model in Equation (2) becomes*

$$Pr(Y_{i1} \leq y_1, Y_{i2} \leq y_2) = C_{\rho_i}^{\text{Normal}}\left\{\Phi\left(\frac{y_1 - \mu_1}{\sigma_{i1}}\right), \Phi\left(\frac{y_2 - \mu_2}{\sigma_{i2}}\right)\right\} = \Phi_{\rho_i}\left(\frac{y_1 - \mu_1}{\sigma_{i1}}, \frac{y_2 - \mu_2}{\sigma_{i2}}\right).$$

Under this model, the correlation function is the identity function $\rho_{C^{\text{Normal}}}(\rho) = \rho$. In addition, one has the copula parameter space $\Theta_{C^{\text{Normal}}} = (-1, 1)$, and the range of correlations $R_{C^{\text{Normal}}} = (-1, 1)$. Without doubt, for any $\rho_i \in (-1, 1)$, the copula parameter can be identified.

Example 2. *(the FGM copula): Under the FGM copula, the model in Equation (2) becomes*

$$Pr(Y_{i1} \leq y_1, Y_{i2} \leq y_2) = \Phi\left(\frac{y_1 - \mu_1}{\sigma_{i1}}\right)\Phi\left(\frac{y_2 - \mu_2}{\sigma_{i2}}\right)\left[1 + \theta_i\left\{1 - \Phi\left(\frac{y_1 - \mu_1}{\sigma_{i1}}\right)\right\}\left\{1 - \Phi\left(\frac{y_2 - \mu_2}{\sigma_{i2}}\right)\right\}\right].$$

Under this model, the correlation function is $\rho_{C^{\text{FGM}}}(\theta) = \theta/\pi$ for $-1 \leq \theta \leq 1$ [44]. Thus, the copula parameter is identified by $\theta_i = \pi\rho_i$, as long as $\rho_i \in [-1/\pi, 1/\pi] \approx [-0.32, 0.32]$. If $\rho_i \notin [-1/\pi, 1/\pi]$, we suggest $\theta_i = -1$ or $\theta_i = 1$, using $\theta_i \equiv \pi\rho_i^*$, where $\rho_i^* = min[-1/\pi, max\{\rho_i, 1/\pi\}]$. Hence, θ_i can still be identified by ρ_i. This boundary enforcement is illustrated in Figure 1.
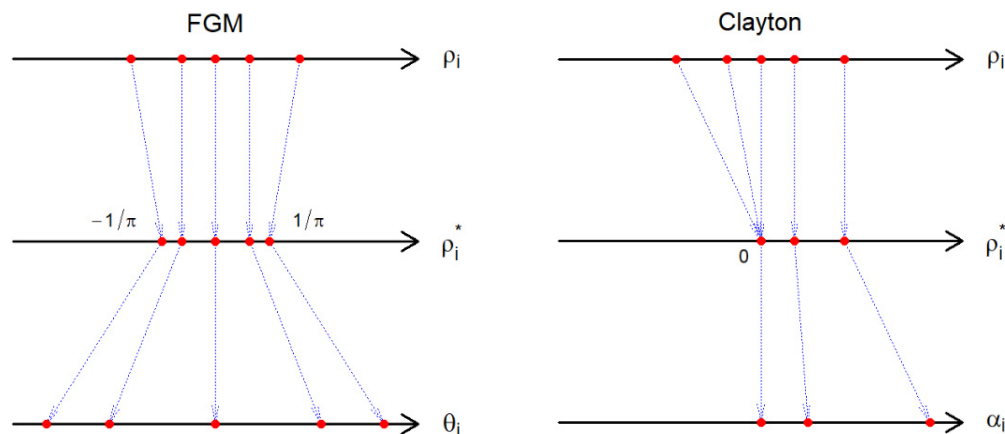

Figure 1. The boundary correction for the correlation coefficient under the bivariate FGM and Clayton models. The first step forces the correlation ρ_i to fall in a range that can be modeled by the chosen copula. The second step transforms the corrected correlation ρ_i^* to the corresponding copula parameter.

Example 3. *(The Clayton copula): Under the Clayton copula, the model in Equation (2) becomes*

$$Pr(Y_{i1} \leq y_1, Y_{i2} \leq y_2) = \left\{\Phi\left(\frac{y_1 - \mu_1}{\sigma_{i1}}\right)^{-\alpha_i} + \Phi\left(\frac{y_2 - \mu_2}{\sigma_{i2}}\right)^{-\alpha_i} - 1\right\}^{-1/\alpha_i}.$$

for $\alpha_i > 0$. The correlation function does not have a closed-form, and is written as

$$\rho_{C^{\text{Clayton}}}(\alpha) = (\alpha + 1) \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \frac{z_1 z_2 \varphi(z_1) \varphi(z_2)}{\Phi(z_1)^{\alpha+1} \Phi(z_2)^{\alpha+1} \left\{ \Phi(z_1)^{-\alpha} + \Phi(z_2)^{-\alpha} - 1 \right\}^{1/\alpha+2}} dz_1 dz_2.$$

It is known that $lim_{\alpha \to 0} \rho_{C^{\text{Clayton}}}(\alpha) = 0$ and $lim_{\alpha \to \infty} \rho_{C^{\text{Clayton}}}(\alpha) = 1$. In addition, if $\alpha_2 \geq \alpha_1$ then $C_{\alpha_2}^{\text{Clayton}}(u,v) \geq C_{\alpha_1}^{\text{Clayton}}(u,v)$ for all $u, v \in (0,1)$ [45]. Then, we conclude that the range of the correlation is $R_{C^{\text{Clayton}}} = (0,1)$. Thus, one can identify by solving $\rho_{C^{\text{Clayton}}}(\alpha_i) = \rho_i$ numerically if $\rho_i > 0$. If $\rho_i \leq 0$, we suggest the independence model (Figure 1)

$$Pr(Y_{i1} \leq y_1, Y_{i2} \leq y_2) = \Phi\left(\frac{y_1 - \mu_1}{\sigma_{i1}}\right) \Phi\left(\frac{y_2 - \mu_2}{\sigma_{i2}}\right).$$

Example 4. *(The Gumbel copula): Under the Gumbel copula, the model in Equation* (2) *becomes*

$$Pr(Y_{i1} \leq y_1, Y_{i2} \leq y_2) = \exp\left(-\left[\left\{-\log \Phi\left(\frac{y_1 - \mu_1}{\sigma_{i1}}\right)\right\}^{\beta_i} + \left\{-\log \Phi\left(\frac{y_2 - \mu_2}{\sigma_{i2}}\right)\right\}^{\beta_i}\right]^{1/\beta_i}\right)$$

for $\beta_i \geq 1$. Similar to the Clayton copula, the correlation function does not have a closed-form, and is not displayed here. It is known that $\rho_{C^{\text{Gumbel}}}(1) = 0$ and $lim_{\beta \to \infty} \rho_{C^{\text{Gumbel}}}(\beta) = 1$. If $\rho_i < 0$, we suggest the independence model as in the Clayton copula.

Example 5. *(The Frank copula): Under the Frank copula, the model in Equation* (2) *becomes*

$$Pr(Y_{i1} \leq y_1, Y_{i2} \leq y_2) = -\frac{1}{\gamma_i} \log\left(1 + \frac{\left[\exp\left\{-\gamma_i \Phi\left(\frac{y_1 - \mu_1}{\sigma_{i1}}\right)\right\} - 1\right]\left[\exp\left\{-\gamma_i \Phi\left(\frac{y_2 - \mu_2}{\sigma_{i2}}\right)\right\} - 1\right]}{e^{-\gamma_i} - 1}\right)$$

for $\gamma_i \neq 0$. Again, the correlation function does not have a closed-form, and is not displayed here. It is known that $lim_{\gamma \to -\infty} \rho_{C^{\text{Frank}}}(\gamma) = -1$ and $lim_{\gamma \to \infty} \rho_{C^{\text{Frank}}}(\gamma) = 1$. Thus, the Frank copula parameter does not require boundary correction.

3.2. Statistical Inference Methods

This subsection develops statistical inference methods under the proposed model.

We propose the MLE for μ under the general copula model (Definition 1) in Equation (2). Suppose that the copula density $C_{\theta}^{[1,1]}$ exists. Then, the joint density of Y_i is

$$f_{i,\mu}(y) = \frac{\partial^2}{\partial y_1 \partial y_2} Pr(Y_{i1} \leq y_1, Y_{i2} \leq y_2) = \frac{1}{\sigma_{i1}\sigma_{i2}} \varphi\left(\frac{y_1 - \mu_1}{\sigma_{i1}}\right) \varphi\left(\frac{y_2 - \mu_2}{\sigma_{i2}}\right) C_{\theta_i}^{[1,1]}\left\{\Phi\left(\frac{Y_1 - \mu_1}{\sigma_{i1}}\right), \Phi\left(\frac{Y_2 - \mu_2}{\sigma_{i2}}\right)\right\}.$$

where $y = (y_1, y_2)$ and $\varphi(\cdot)$ is the density of $N(0,1)$. Given the samples, the log-likelihood function is

$$\ell_n(\mu) = \text{constant} + \sum_{i=1}^{n} \log\left[C_{\theta_i}^{[1,1]}\left\{\Phi\left(\frac{Y_{i1} - \mu_1}{\sigma_{i1}}\right), \Phi\left(\frac{Y_{i2} - \mu_2}{\sigma_{i2}}\right)\right\}\right] - \frac{1}{2} \sum_{j=1}^{2} \sum_{i=1}^{n} \left(\frac{Y_{ij} - \mu_j}{\sigma_{ij}}\right)^2.$$

The MLE of the common mean vector is defined as

$$\hat{\mu}_n = \begin{bmatrix} \hat{\mu}_{n,1} \\ \hat{\mu}_{n,2} \end{bmatrix} = \underset{\mu \in R^2}{\text{argmax}} \ell_n(\mu),$$

where $R = (-\infty, \infty)$ is a real line. The MLE does not have a closed-form expression except for the normal copula. Thus, the MLE can also be obtained by the Newton–Raphson algorithm or some software functions (e.g., the R functions *optim* or *nlm*). One may also apply our R package *CommonMean.Copula* [26], which will be explained in Section 6.

3.3. Information Matrix

For the MLE to be well-behaved, it is necessary to show that the (Fisher) information matrix exists and is non-singular. In other words, the MLE, without verifying these conditions, may have some problems, e.g., the non-existence, inconsistency, or inefficiency of the MLE. Furthermore, the information matrix describes how a copula influences the MLE.

We define the 2×2 information matrix $I_i(\boldsymbol{\mu})$ for $i = 1, 2, \ldots, n$ as

$$I_{i,jk}(\boldsymbol{\mu}) = E\left\{ \frac{\partial \log f_{i,\boldsymbol{\mu}}(\boldsymbol{Y}_i)}{\partial \mu_j} \frac{\partial \log f_{i,\boldsymbol{\mu}}(\boldsymbol{Y}_i)}{\partial \mu_k} \right\}, \quad j, k = 1, 2.$$

The following theorem gives the formula of the information matrix.

Lemma 1. *If $C_{\boldsymbol{\theta}}^{[3,1]}$, $C_{\boldsymbol{\theta}}^{[1,3]}$, and $C_{\boldsymbol{\theta}}^{[2,2]}$ exist in $(0, 1)^2$, for each i, the following equalities hold*

$$E\left\{ \frac{\partial \log f_{i,\boldsymbol{\mu}}(\boldsymbol{Y}_i)}{\partial \mu_j} \frac{\partial \log f_{i,\boldsymbol{\mu}}(\boldsymbol{Y}_i)}{\partial \mu_k} \right\} = E\left\{ -\frac{\partial^2 \log f_{i,\boldsymbol{\mu}}(\boldsymbol{Y}_i)}{\partial \mu_j \partial \mu_k} \right\}; \quad j, k = 1, 2.$$

The proof of Lemma 1 is given in Appendix A.1.

Many copulas have $C_{\boldsymbol{\theta}}^{[3,1]}$, $C_{\boldsymbol{\theta}}^{[1,3]}$, and $C_{\boldsymbol{\theta}}^{[2,2]}$ in $(0, 1)^2$, such as the normal, FGM, and Clayton copulas (Appendix A.2.). The following theorem gives the formula of the information matrix.

Theorem 1. *Under the copula-based model (Definition 1), the information matrix does not depend on $\boldsymbol{\mu}$. Furthermore, if $C_{\boldsymbol{\theta}}^{[3,1]}$, $C_{\boldsymbol{\theta}}^{[1,3]}$, and $C_{\boldsymbol{\theta}}^{[2,2]}$ exist in $(0, 1)^2$, it can be decomposed into the sum of the information matrix for the independent model and the additional information by the copula,*

$$I_i = \begin{bmatrix} \frac{1}{\sigma_{i1}^2} & 0 \\ 0 & \frac{1}{\sigma_{i2}^2} \end{bmatrix} + \begin{bmatrix} \frac{1}{\sigma_{i1}^2} E_C^{11}(\boldsymbol{\theta}_i) & \frac{1}{\sigma_{i1}\sigma_{i2}}\left\{ E_C^{12}(\boldsymbol{\theta}_i) - \rho_C(\boldsymbol{\theta}_i) \right\} \\ \frac{1}{\sigma_{i1}\sigma_{i2}}\left\{ E_C^{12}(\boldsymbol{\theta}_i) - \rho_C(\boldsymbol{\theta}_i) \right\} & \frac{1}{\sigma_{i2}^2} E_C^{22}(\boldsymbol{\theta}_i) \end{bmatrix}, \quad (3)$$

where

$$E_C^{11}(\boldsymbol{\theta}_i) = E\left\{ \frac{\varphi(Z_{i1}) C_{\boldsymbol{\theta}_i}^{[2,1]}\{\Phi(Z_{i1}), \Phi(Z_{i2})\}}{C_{\boldsymbol{\theta}_i}^{[1,1]}\{\Phi(Z_{i1}), \Phi(Z_{i2})\}} \right\}^2,$$

$$E_C^{22}(\boldsymbol{\theta}_i) = E\left\{ \frac{\varphi(Z_{i2}) C_{\boldsymbol{\theta}_i}^{[1,2]}\{\Phi(Z_{i1}), \Phi(Z_{i2})\}}{C_{\boldsymbol{\theta}_i}^{[1,1]}\{\Phi(Z_{i1}), \Phi(Z_{i2})\}} \right\}^2,$$

$$E_C^{12}(\boldsymbol{\theta}_i) = E\left[\frac{\varphi(Z_{i1}) \varphi(Z_{i2}) C_{\boldsymbol{\theta}_i}^{[2,1]}\{\Phi(Z_{i1}), \Phi(Z_{i2})\} C_{\boldsymbol{\theta}_i}^{[1,2]}\{\Phi(Z_{i1}), \Phi(Z_{i2})\}}{C_{\boldsymbol{\theta}_i}^{[1,1]}\{\Phi(Z_{i1}), \Phi(Z_{i2})\}^2} \right].$$

Theorem 1 can be proved by straightforward calculations as Lemma 1 (Appendix A.1.). Theorem 1 helps us interpret the role of the copula $C_{\boldsymbol{\theta}_i}$ on the information matrix.

Theorem 2. *The determinant of I_i can be expressed as*

$$\det(I_i) = \frac{1}{\sigma_{i1}^2 \sigma_{i2}^2}\left\{ E_C^{11}(\boldsymbol{\theta}_i) E_C^{22}(\boldsymbol{\theta}_i) - E_C^{12}(\boldsymbol{\theta}_i)^2 \right\}$$

$$+ \frac{1}{\sigma_{i1}^2 \sigma_{i2}^2}\left\{ E_C^{11}(\boldsymbol{\theta}_i) + E_C^{22}(\boldsymbol{\theta}_i) + 2\rho_C(\boldsymbol{\theta}_i) E_C^{12}(\boldsymbol{\theta}_i) \right\} + \frac{1}{\sigma_{i1}^2 \sigma_{i2}^2}\left\{ 1 - \rho_C(\boldsymbol{\theta}_i)^2 \right\}.$$

In addition, $\det(I_i) > 0$ and I_i is positive definite.

Proof of Theorem 2. The expression of $\det(I_i)$ is obtained by straightforward calculations. Clearly, we have $|\rho_C(\boldsymbol{\theta}_i)| < 1$. Then, by the Cauchy-Schwarz inequality,

$$E_C^{11}(\boldsymbol{\theta}_i)E_C^{22}(\boldsymbol{\theta}_i) \geq E_C^{12}(\boldsymbol{\theta}_i)^2.$$

Furthermore, by the arithmetic-geometric mean inequality, we have

$$E_C^{11}(\boldsymbol{\theta}_i) + E_C^{22}(\boldsymbol{\theta}_i) \geq 2\left\{E_C^{11}(\boldsymbol{\theta}_i)E_C^{22}(\boldsymbol{\theta}_i)\right\}^{1/2} \geq 2\left|E_C^{12}(\boldsymbol{\theta}_i)\right| > 2\left|\rho_C(\boldsymbol{\theta}_i)E_C^{12}(\boldsymbol{\theta}_i)\right|.$$

Then we obtain $\det(I_i) > 0$. Since $E_C^{11}(\boldsymbol{\theta}_i)/\sigma_{i1}^2 + 1/\sigma_{i1}^2 > 0$, both the upper left 1×1 and 2×2 determinants of I_i are positive. Thus, I_i is positive definite. $\square$

Based on Theorem 1, one can derive the information matrix $I_i(\boldsymbol{\mu})$ for parametric copulas. Below, we show examples of the normal, FGM, and Clayton copulas.

Example 6. *(The normal copula): Under the normal copula*

$$E_{C\text{Normal}}^{11}(\rho_i) = E_{C\text{Normal}}^{22}(\rho_i) = \frac{\rho_i^2}{1 - \rho_i^2}, \quad E_{C\text{Normal}}^{12}(\rho_i) = -\frac{\rho_i^3}{1 - \rho_i^2}.$$

Then, by Theorem 1, the information matrix in Equation (3) becomes

$$I_i = \begin{bmatrix} \frac{1}{\sigma_{i1}^2} & 0 \\ 0 & \frac{1}{\sigma_{i2}^2} \end{bmatrix} + \begin{bmatrix} \frac{1}{\sigma_{i1}^2}\frac{\rho_i^2}{1-\rho_i^2} & \frac{1}{\sigma_{i1}\sigma_{i2}}\left(-\frac{\rho_i^3}{1-\rho_i^2} - \rho_i\right) \\ \frac{1}{\sigma_{i1}\sigma_{i2}}\left(-\frac{\rho_i^3}{1-\rho_i^2} - \rho_i\right) & \frac{1}{\sigma_{i2}^2}\frac{\rho_i^2}{1-\rho_i^2} \end{bmatrix}$$

$$= \frac{1}{1-\rho_i^2}\begin{bmatrix} \frac{1}{\sigma_{i1}^2} & -\frac{\rho_i}{\sigma_{i1}\sigma_{i2}} \\ -\frac{\rho_i}{\sigma_{i1}\sigma_{i2}} & \frac{1}{\sigma_{i2}^2} \end{bmatrix} = \boldsymbol{\Omega}_i^{-1}.$$

and its determinant is $\det(I_i^{\text{Normal}}) = 1/(\sigma_{i1}^2\sigma_{i2}^2)$. *Clearly,* I_i^{Normal} *is positive definite.*

Example 7. *(The FGM copula): Under the FGM copula*

$$E_{C\text{FGM}}^{11}(\theta_i) = E_{C\text{FGM}}^{22}(\theta_i) = 4\theta_i^2 \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \frac{\varphi(z_1)^3\{1 - 2\Phi(z_2)\}^2\varphi(z_2)}{1 + \theta_i\{1 - 2\Phi(z_1)\}\{1 - 2\Phi(z_2)\}}dz_1dz_2,$$

$$E_{C\text{FGM}}^{12}(\theta_i) = 4\theta_i^2 \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \frac{\varphi(z_1)^2\{1 - 2\Phi(z_1)\}\varphi(z_2)^2\{1 - 2\Phi(z_2)\}}{1 + \theta_i\{1 - 2\Phi(z_1)\}\{1 - 2\Phi(z_2)\}}dz_1dz_2, \quad \rho_{C\text{FGM}}(\theta) = \frac{\theta}{\pi}.$$

Then, by Theorem 1, the information matrix in Equation (3) becomes

$$I_i^{\text{FGM}} = \begin{bmatrix} \frac{1}{\sigma_{i1}^2} & 0 \\ 0 & \frac{1}{\sigma_{i2}^2} \end{bmatrix} + \begin{bmatrix} \frac{1}{\sigma_{i1}^2}E_{C\text{FGM}}^{11}(\theta_i) & \frac{1}{\sigma_{i1}\sigma_{i2}}E_{C\text{FGM}}^{12}(\theta_i) - \frac{\theta_i}{\pi\sigma_{i1}\sigma_{i2}} \\ \frac{1}{\sigma_{i1}\sigma_{i2}}E_{C\text{FGM}}^{12}(\theta_i) - \frac{\theta_i}{\pi\sigma_{i1}\sigma_{i2}} & \frac{1}{\sigma_{i2}^2}E_{C\text{FGM}}^{22}(\theta_i) \end{bmatrix}.$$

By Theorem 2, its determinant is

$$\det(I_i) = \frac{1}{\sigma_{i1}^2\sigma_{i2}^2}\left\{E_{C\text{FGM}}^{11}(\theta_i)^2 - E_{C\text{FGM}}^{12}(\theta_i)^2\right\}$$

$$+ \frac{2}{\sigma_{i1}^2\sigma_{i2}^2}\left\{E_{C\text{FGM}}^{11}(\theta_i) + \frac{\theta_i}{\pi}E_{C\text{FGM}}^{12}(\theta_i)\right\} + \frac{1}{\sigma_{i1}^2\sigma_{i2}^2}\left(1 - \frac{\theta_i^2}{\pi^2}\right).$$

This result agrees with [3] who considered the FGM model.

Example 8. *(The Clayton copula): Under the Clayton copula*

$$E^{11}_{C^{\text{Clayton}}}(\alpha_i) = E^{22}_{C^{\text{Clayton}}}(\alpha_i) = (\alpha_i+1) \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \frac{\varphi(z_1)^3 \varphi(z_2)\left\{ \alpha_i \Phi(z_1)^{-\alpha_i} - (\alpha_i+1)\Phi(z_2)^{-\alpha_i} + (\alpha_i+1) \right\}^2}{\Phi(z_1)^{\alpha_i+3}\Phi(z_2)^{\alpha_i+1}\left\{ \Phi(z_1)^{-\alpha_i} + \Phi(z_2)^{-\alpha_i} - 1 \right\}^{1/\alpha_i+4}} dz_1 dz_2,$$

$$E^{12}_{C^{\text{Clayton}}}(\alpha_i) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \frac{\varphi(z_1)^2\left\{ \alpha_i \Phi(z_1)^{-\alpha_i} - (\alpha_i+1)\Phi(z_2)^{-\alpha_i} + (\alpha_i+1) \right\}}{\Phi(z_1)^{\alpha_i+2}\left\{ \Phi(z_1)^{-\alpha_i} + \Phi(z_2)^{-\alpha_i} - 1 \right\}^{1/2\alpha_i+2}}$$
$$\times \frac{\varphi(z_2)^2\left\{ \alpha_i \Phi(z_2)^{-\alpha_i} - (\alpha_i+1)\Phi(z_1)^{-\alpha_i} + (\alpha_i+1) \right\}}{\Phi(z_2)^{\alpha_i+2}\left\{ \Phi(z_1)^{-\alpha_i} + \Phi(z_2)^{-\alpha_i} - 1 \right\}^{1/2\alpha_i+2}} dz_1 dz_2,$$

$$\rho_{C^{\text{Clayton}}}(\alpha_i) = (\alpha_i+1)\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \frac{z_1 z_2 \varphi(z_1)\varphi(z_2)}{\Phi(z_1)^{\alpha_i+1}\Phi(z_2)^{\alpha_i+1}\left\{ \Phi(z_1)^{-\alpha_i} + \Phi(z_2)^{-\alpha_i} - 1 \right\}^{1/\alpha_i+2}} dz_1 dz_2.$$

Then, by Theorems 1 and 2, we obtain I_i^{Clayton} and $\det(I_i^{\text{Clayton}})$ accordingly.

4. Asymptotic Theory

To assess the sampling variability of $\hat{\boldsymbol{\mu}}_n$, its asymptotic distribution is presented in this section.

A technical burden comes from the fact that our samples $\boldsymbol{Y}_i$, $i = 1,2,\ldots,n$ are independent and *non-identically* distributed (i.n.i.d.) owing to heterogeneous variances ($\boldsymbol{\Omega}_i \neq \boldsymbol{\Omega}_j, i \neq j$). The existence of the asymptotic distribution requires the stabilization of the information matrix [3,46,47] in large samples. For the asymptotic variance of $\hat{\boldsymbol{\mu}}_n$, to be defined, we assume the existence of a 2×2 positive definite matrix $I \equiv \lim_{n\to\infty} \sum_{i=1}^{n} I_i/n$. We further assume that the copula's derivatives $C_{\theta}^{[4,1]}$, $C_{\theta}^{[3,2]}$, $C_{\theta}^{[2,3]}$, and $C_{\theta}^{[1,4]}$ exist in $(0,1)^2$. With these conditions and many other technical conditions given in [48], we establish the consistency and asymptotic normality of $\hat{\boldsymbol{\mu}}_n$:

Theorem 3. *Under the copula model (Definition 1), if some regularity conditions hold, then*

(a)   *Existence and consistency: With probability tending to one, there exists the MLE $\hat{\boldsymbol{\mu}}_n = (\hat{\mu}_{n,1}, \hat{\mu}_{n,2})$ such that $\hat{\boldsymbol{\mu}}_n \to {}_p\boldsymbol{\mu}$, as $n \to \infty$;*

(b)   *Asymptotic normality: $n^{1/2}(\hat{\boldsymbol{\mu}}_n - \boldsymbol{\mu}) \to {}_d N(\boldsymbol{0}, I^{-1})$, as $n \to \infty$.*

The proof of Theorem 3 and the required regularity conditions are given in the Ph.D dissertation of [48]. The proof approximates $n^{1/2}(\hat{\boldsymbol{\mu}}_n - \boldsymbol{\mu})$ by the sum of independent random variables, and then applies the weak law of large numbers for i.n.i.d. random variables from Theorem 1.14 in [49] and the Lindeberg–Feller multivariate central limit theorem from Proposition 2.27 in [50]. The proof is fairly technical, but similar to those of Theorem 6.5.1 in [51], Theorem 1 in [47], and Theorem 5.1 in [3].

5. SE and Confidence Sets

As Section 4 has established the asymptotic theory to evaluate the variability of the proposed MLE, we can derive the SE, confidence interval (CI), and confidence ellipse (CE).

Let $g : R^2 \mapsto R$ be a differentiable function, and $g(\boldsymbol{\mu})$ be the parameter of interest. For instance, $g(\boldsymbol{\mu}) = \mu_1$ and $g(\boldsymbol{\mu}) = \mu_2 - \mu_1$ can be considered. The SE of $g(\hat{\boldsymbol{\mu}}_n)$ is

$$\text{SE}\{g(\hat{\boldsymbol{\mu}}_n)\} = \left[\left\{ \frac{\partial g(\boldsymbol{\mu})}{\partial \boldsymbol{\mu}} \right\}^T \left\{ -\frac{\partial^2 \ell_n(\boldsymbol{\mu})}{\partial \boldsymbol{\mu} \partial \boldsymbol{\mu}^T} \right\}^{-1} \left\{ \frac{\partial g(\boldsymbol{\mu})}{\partial \boldsymbol{\mu}} \right\} \Bigg|_{\boldsymbol{\mu} = \hat{\boldsymbol{\mu}}_n} \right]^{1/2}.$$

This formula is based on the delta method and the large sample approximation

$$I \approx \frac{1}{n} \sum_{i=1}^{n} I_i \approx -\frac{1}{n} \frac{\partial^2 \ell_n(\boldsymbol{\mu})}{\partial \boldsymbol{\mu} \partial \boldsymbol{\mu}^T} \bigg|_{\boldsymbol{\mu} = \hat{\boldsymbol{\mu}}_n}.$$

The 95% CI for $g(\boldsymbol{\mu})$ is $g(\hat{\boldsymbol{\mu}}_n) \pm 1.96 \times \text{SE}\{g(\hat{\boldsymbol{\mu}}_n)\}$.

Moreover, based on Theorem 3, we construct a 95% CE for $\boldsymbol{\mu}$:

$$\text{CE} = \left\{ \boldsymbol{\mu} : (\hat{\boldsymbol{\mu}}_n - \boldsymbol{\mu})^T \left(-\frac{\partial^2 \ell_n(\boldsymbol{\mu})}{\partial \boldsymbol{\mu} \partial \boldsymbol{\mu}^T} \bigg|_{\boldsymbol{\mu} = \hat{\boldsymbol{\mu}}_n} \right) (\hat{\boldsymbol{\mu}}_n - \boldsymbol{\mu}) \leq \chi^2_{\text{df}=2,0.95} \right\},$$

where $\chi^2_{\text{df}=2,0.95}$ is be the 95% point of the χ^2-distribution with two degrees of freedom.

6. R Package

We implement the proposed methods in an R package *CommonMean.Copula* [26]. R users can easily compute the MLE with its SE and 95% CI under the FGM, Clayton, Gumbel, Frank, and normal copulas. In this package, the log-likelihood is maximized by the R *optim* function, where the initial values are set as the univariate estimators

$$\mu_j^{(0)} = \left(\sum_{i=1}^{n} \frac{1}{\sigma_{ij}^2} \right)^{-1} \sum_{i=1}^{n} \frac{Y_{ij}}{\sigma_{ij}^2}, \qquad j = 1, 2.$$

For illustration, we fitted the Clayton copula by the following R codes:

```
> library(CommonMean.Copula)
> Y1 = c(35,25,30,50,60) # outcome 1
> Y2 = c(30,30,50,65,40) # outcome 2
> Sigma1 = c(1.3,1.4,1.5,2.0,1.8) # SE of outcome 1
> Sigma2 = c(1.7,1.9,2.5,2.2,1.8) # SE of outcome 2
> rho = c(0.4,0.7,0.6,0.7,0.6) # correlation between two outcomes
> CommonMean.Copula(Y1,Y2,Sigma1,Sigma2,rho,copula = "clayton") # input
```

- some outputs are omitted for brevity –

```
$`CommonMean 1`
  estimate         SE      Lower      Upper
33.9505331  0.4393516 33.0894198 34.8116463

$`CommonMean 2`
  estimate         SE      Lower      Upper
41.9926717  0.6039963 40.8088607 43.1764827

$v
          [,1]       [,2]
[1,] 0.1930298 0.2507297
[2,] 0.2507297 0.3648115

$`Log-likelihood values`
          1          2          3          4          5      total
 -46.30760  -21.14432  -18.78560  -89.05592 -110.36095 -285.65438
```

Here, $CommonMean1 shows $\hat{\mu}_{n,1}^{\text{Clayton}} = 33.95$, $SE(\hat{\mu}_{n,1}^{\text{Clayton}}) = 0.439$, and the 95% CI (33.089, 34.812); $CommonMean2 is similar. $V shows the covariance matrix $Cov(\hat{\boldsymbol{\mu}}_n^{\text{Clayton}})$. $`Log-likelihood values` shows $\ell_n^{\text{Clayton}}(\hat{\boldsymbol{\mu}}_n^{\text{Clayton}}) = -285.65$. One can fit other copulas by changing "Clayton" to "FGM", "Gumbel", "Frank", or "normal".

7. Simulation Studies

We conducted Monte Carlo simulations to examine the accuracy of the proposed methods. We report the results for the Clayton copula; more results are available from [48].

We generated Y_i, $i = 1, 2, \ldots, n$, under the Clayton copula with $\alpha_i \sim Gamma(64, 1/8)$, $Gamma(4, 1/2)$, or $Gamma(1, 1)$, leading to $E[\alpha_i] = 8$, $E[\alpha_i] = 2$, or $E[\alpha_i] = 1$, respectively. In all three cases, we have $Var[\alpha_i] = 1$. Without loss of generality, we set $\boldsymbol{\mu} = (0, 0)$. To set σ_{i1}^2 and σ_{i2}^2, we followed the simulation setting of [52]. That is, $\sigma_{i1}^2, \sigma_{i2}^2 \sim \chi_{df=1}^2/4$, restricted in the interval [0.009, 0.6]. This setting leads to $E[\sigma_{i1}^2] = E[\sigma_{i2}^2] = 0.173$. Based on the generated data, we computed $\hat{\mu}_{n,1}^{\text{Clayton}}$, $\hat{\mu}_{n,2}^{\text{Clayton}} - \hat{\mu}_{n,1}^{\text{Clayton}}$, and $\hat{\boldsymbol{\mu}}_n^{\text{Clayton}}$, and their SEs and 95% CIs (CEs) by using the R function *CommonMean.Copula* (Section 6). We then evaluated the coverage probability (CP) of the 95% CI (CE) to see how the confidence set can cover the true value. We consider a small sample size $n \in \{5, 10, 15\}$ and a large sample size $n \in \{50, 100, 300\}$. Our simulations are based on 1000 repetitions.

Table 1 summarizes the results. For $\hat{\mu}_{n,1}^{\text{Clayton}}$ and $\hat{\mu}_{n,2}^{\text{Clayton}} - \hat{\mu}_{n,1}^{\text{Clayton}}$, the SDs of the estimates decrease when n increases from $n = 5$ to $n = 300$. We report the boxplots summarizing the 1000 repetitions for $\hat{\mu}_{n,1}^{\text{Clayton}}$ in Figure 2. This clearly visualizes how the variability of the estimates vanishes as the sample sizes increase. Table 1 also shows that the SDs are close to the average SEs, except for $n = 5$ (due to the very small samples). Consequently, the CPs are close enough to the nominal level of 0.95, especially when sample sizes are large, which is consistent with our asymptotic theories. For $\hat{\boldsymbol{\mu}}_n^{\text{Clayton}}$, the CPs of the 95% CEs are also reasonably close to the nominal level. In summary, the proposed estimators and the asymptotic theory work fairly well in finite samples.

Table 1. Simulation results based on 1000 repetitions.

		$\hat{\mu}_{n,1}^{\text{Clayton}}$			$\hat{\mu}_{n,2}^{\text{Clayton}} - \hat{\mu}_{n,1}^{\text{Clayton}}$			$\hat{\boldsymbol{\mu}}_n^{\text{Clayton}}$
Parameters	n	SD	SE	CP	SD	SE	CP	CP
$E[\alpha_i] = 8$	5	0.064	0.046	0.888	0.042	0.033	0.885	0.859
	10	0.033	0.026	0.913	0.023	0.019	0.931	0.894
	15	0.021	0.019	0.933	0.015	0.014	0.936	0.919
	50	0.010	0.009	0.952	0.007	0.007	0.954	0.950
	100	0.007	0.006	0.955	0.005	0.005	0.943	0.944
	300	0.004	0.004	0.948	0.003	0.003	0.942	0.948
$E[\alpha_i] = 2$	5	0.105	0.092	0.938	0.100	0.086	0.920	0.909
	10	0.061	0.057	0.937	0.060	0.055	0.929	0.919
	15	0.049	0.045	0.938	0.045	0.042	0.943	0.930
	50	0.023	0.023	0.959	0.022	0.021	0.944	0.943
	100	0.016	0.016	0.942	0.015	0.015	0.957	0.950
	300	0.009	0.009	0.946	0.008	0.008	0.946	0.949
$E[\alpha_i] = 1$	5	0.115	0.105	0.932	0.128	0.120	0.935	0.922
	10	0.069	0.068	0.950	0.079	0.075	0.937	0.937
	15	0.058	0.053	0.941	0.062	0.058	0.947	0.937
	50	0.028	0.027	0.942	0.029	0.030	0.955	0.943
	100	0.019	0.019	0.942	0.021	0.020	0.936	0.945
	300	0.010	0.011	0.958	0.011	0.012	0.960	0.959

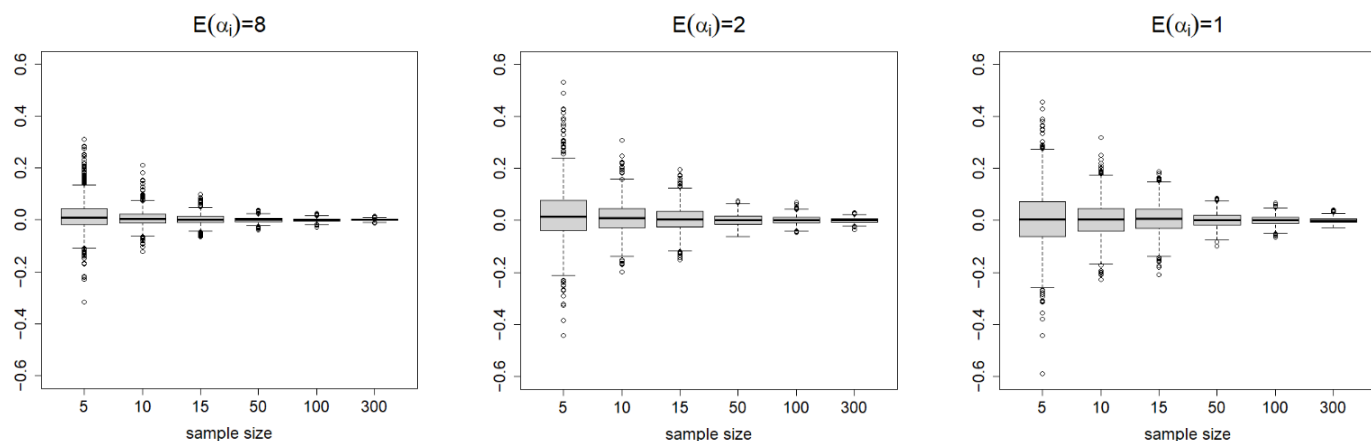SD = standard deviation, SE = standard error, CP = coverage probability of the 95% CI (CE).

Figure 2. The boxplots summarizing the 1000 repetitions for $\hat{\mu}_{n,1}^{\text{Clayton}}$ with true parameter $\mu_1 = 0$ under the copula parameters $E[\alpha_i] = 8$, $E[\alpha_i] = 2$, or $E[\alpha_i] = 1$. The sample size varies from $n = 5$ to $n = 300$.

8. Data Analysis

We analyze two real datasets to illustrate the usefulness of the proposed methods.

8.1. The Entrance Exam Data

The first dataset we analyzed was the entrance exam scores on mathematics and statistics, which was introduced by [3]. The data come from undergrad students who took written exams from 2013 to 2017 to enter the Graduate Institute of Statistics, National Central University, Taiwan. The possible score range is from 0 to 100 for both subjects. Let $i = 1, 2, \ldots, 5$ be indices for years 2013, 2014, $\ldots$, 2017. Table 2 provides the data, including the values of mathematics (Y_{i1} = mean math score) and statistics (Y_{i2} = mean stat score), and their covariance matrix (Ω_i).

Table 2. The entrance exam data from [3].

i	Year	Mean Math Score (Y_{i1})	Mean Stat Score (Y_{i2})	Covariance Matrix (Ω_i)	Copula Parameter				
					ρ_i	θ_i	α_i	β_i	γ_i
1	2013	35.17	30.41	1.77 0.89 0.89 2.99	0.38	1.00	0.67	1.34	2.68
2	2014	23.43	31.63	1.89 1.76 1.76 3.61	0.67	1.00	1.92	1.90	6.00
3	2015	30.74	48.11	2.15 2.12 2.12 6.13	0.58	1.00	1.37	1.65	4.67
4	2016	50.91	65.22	3.87 2.91 2.91 5.02	0.66	1.00	1.82	1.85	5.76
5	2017	61.62	40.22	3.17 2.10 2.10 3.29	0.65	1.00	1.75	1.83	5.60

ρ_i = the Pearson correlation; θ_i = the FGM copula parameter; α_i = the Clayton copula parameter; β_i = the Gumbel copula parameter; γ_i = the Frank copula parameter.

We fitted the data to the proposed model using the R function *CommonMean.Copula(.)* in our R package (Section 6). Table 3 summarizes the fitted results for the FGM, Clayton, Gumbel, Frank, and normal copulas. According to the values of the log-likelihood, the Gumbel copula produces the best fit, the Frank copula the second best, and the bivariate normal model the worst fit. The FGM copula failed to capture the dependence and fitted at the boundary $\theta_i = 1$ for all i (Table 2).

Table 3. Estimation results for the entrance exam data.

Copula	Math: Estimate (95% CI)	Stat: Estimate (95% CI)	Log-likelihood	CV
FGM	37.16 (35.85, 38.47)	41.17 (39.65, 42.70)	−291.80	2723.91
Clayton	32.56 (31.71, 33.40)	43.80 (42.55, 45.05)	−322.84	2644.03
Gumbel	37.67 (36.30, 39.03)	42.56 (40.79, 44.33)	−279.28	2860.21
Frank	37.23 (35.97, 38.49)	39.76 (38.13, 41.40)	−287.63	2738.09
Normal	35.83 (34.51, 37.16)	38.64 (36.94, 40.34)	−342.65	2773.41

Since the number of unknown parameters across different copulas is the same, model selection by the Akaike information criterion (AIC) is equivalent to model selection by the log-likelihood value. An alternative way of selecting a copula is based on a leave-one-out cross validation (*CV*), defined as

$$CV = \sum_{i=1}^{n} \left\{ \left(Y_{i1} - \hat{\mu}_{n,1}^{(-i)} \right)^2 + \left(Y_{i2} - \hat{\mu}_{n,2}^{(-i)} \right)^2 \right\},$$

where $\hat{\mu}_{n,1}^{(-i)}$ and $\hat{\mu}_{n,2}^{(-i)}$ are the MLE obtained without the *i*th sample. Here, *CV* measures how a sample is predicted by the others under a copula model. A smaller *CV* corresponds to a better performance of the model.

Table 3 reports the values of *CV* for each copula. It shows that the Clayton copula has the best performance while the Gumbel copula has the worst. The normal copula has the second worst performance. Overall, our analysis clearly shows the insufficiency of the bivariate normal model.

Figure 3 shows the 95% CEs for the mean vector *μ*. This visualizes how the resultant estimates vary from the choice of copulas. Interestingly, the CE under the Clayton copula is far away from the other four, although it has a larger log-likelihood value than the normal copula. The normal and Clayton copulas produce the rotated oval shape of the CEs, representing a positive dependence between math and stat scores. The FGM and Gumbel copulas produce similar shapes for their CE. We adopt the 95% CE given by the Gumbel copula because it has the largest log-likelihood value.

Figure 4 gives the 3D plot of the log-likelihood surface under the Gumbel copula model. The plot shows that the estimate of the common mean $\hat{\mu}_n^{\text{Gumbel}} = (37.67, 42.56)$ attains the global maximum of the log-likelihood function.
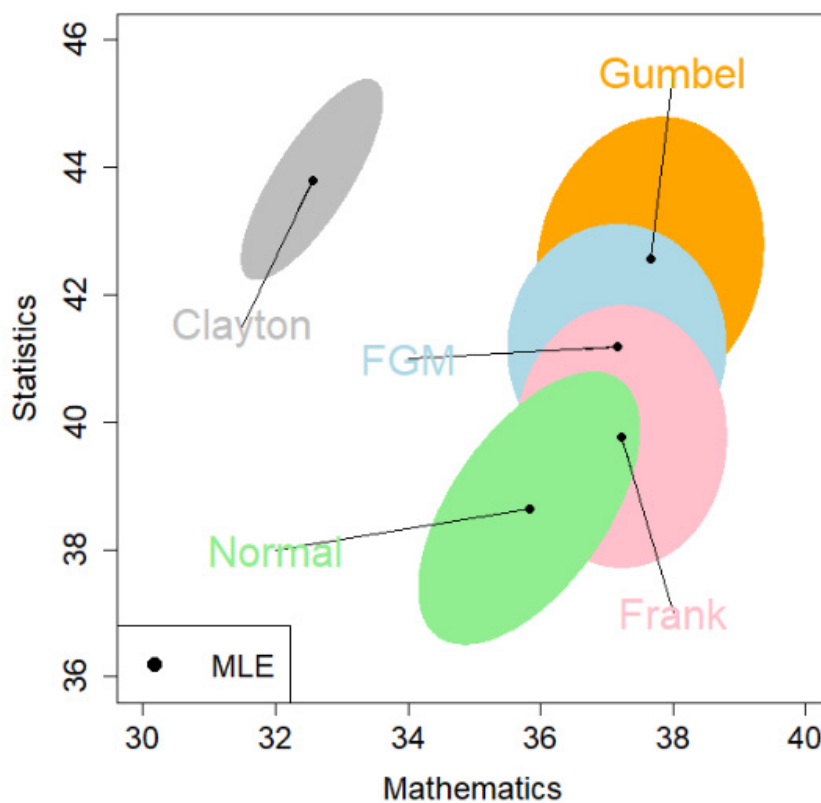
Figure 3. The MLE and the 95% CE (colored region) for the common means based on the exam score data. The copulas are signified by colors: blue = FGM; gray = Clayton; orange = Gumbel; pink = Frank; green = normal.
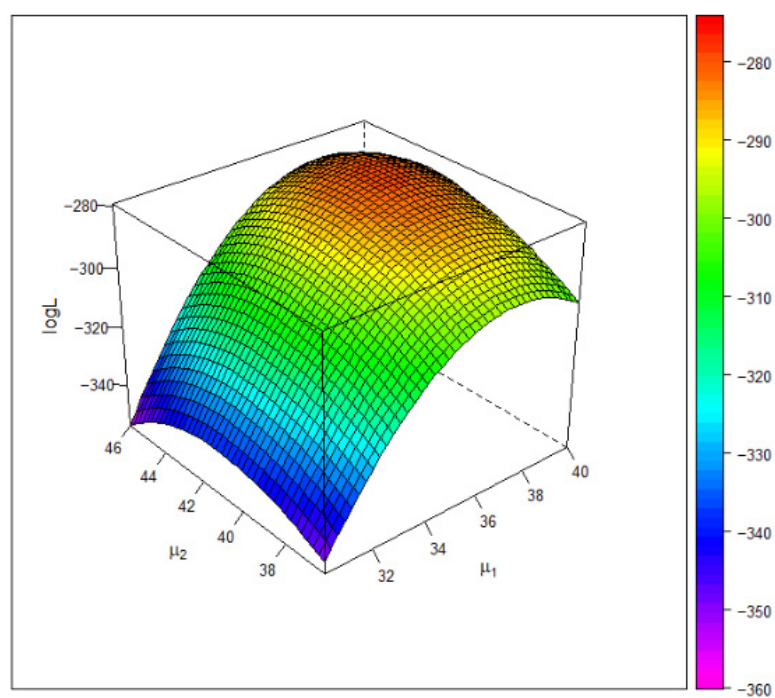

Figure 4. The 3D plot of the log-likelihood value under the Gumbel copula based on the entrance exam data. The maximum occurs at $\hat{\boldsymbol{\mu}}_n^{\text{Gumbel}} = (37.67, 42.56)$.

8.2. The Blood Pressure Data

The second dataset we used contains 10 studies that examined the effectiveness of hypertension treatment for lowering blood pressure. Each study provides complete data on two treatment effects, the difference in systolic blood pressure (SBP) and diastolic blood pressure (DBP) between the treatment and the control groups, where these differences are adjusted for the participants' baseline blood pressures. The within-study correlations of the two outcomes range from $\rho_i = 0.45$ to $\rho_i = 0.78$, exhibiting positive dependence. This dataset is available in R package *mvmeta* [30] and was previously analyzed by [53].

We fitted the data to the proposed copula models using the R function *Common-Mean.Copula(.)* in our R package (Section 6). Table 4 summarizes the fitted results for all the copulas. Based on the log-likelihood values, the Frank copula produces the best fit, the Gumbel copula the second best, and the Clayton copula produces the worst fit. The FGM copula failed to capture the dependence and fitted at the boundary $\theta_i = 1$ for all i. Again, our analysis reveals the insufficiency of the bivariate normal model; the Frank copula best captured the correlations in the blood pressure data. We also compared *CV* across all the copulas (Table 4). The results show that the Clayton copula has the best performance while the normal copula has the worst. Again, our analysis shows the insufficiency of the normal model.

Table 4. Estimation results for the blood pressure data.

Copula	SBP: Estimate (95% CI)	DBP: Estimate (95% CI)	Log-likelihood	*CV*
FGM	$-9.18\,(-9.32, -9.04)$	$-3.94\,(-4.00, -3.89)$	-530.29	177.23
Clayton	$-9.53\,(-9.70, -9.36)$	$-4.34\,(-4.38, -4.29)$	-787.02	163.04
Gumbel	$-8.99\,(-9.16, -8.83)$	$-3.94\,(-4.00, -3.89)$	-514.67	184.67
Frank	$-9.20\,(-9.40, -9.00)$	$-3.94\,(-3.99, -3.88)$	-513.34	179.81
Normal	$-8.43\,(-8.60, -8.25)$	$-3.95\,(-4.01, -3.90)$	-771.82	206.49

SBP = the difference in systolic blood pressure; DBP = the difference in diastolic blood pressure.

Figure 5 shows the 95% CEs for the mean vector $\boldsymbol{\mu}$. The CE under the Clayton and normal copula are far away from the other three. The CE under the FGM copula was almost fully covered by the CE under the Frank copula. We adopt the 95% CE given by the Frank copula, since it has the largest log-likelihood value (Table 4).
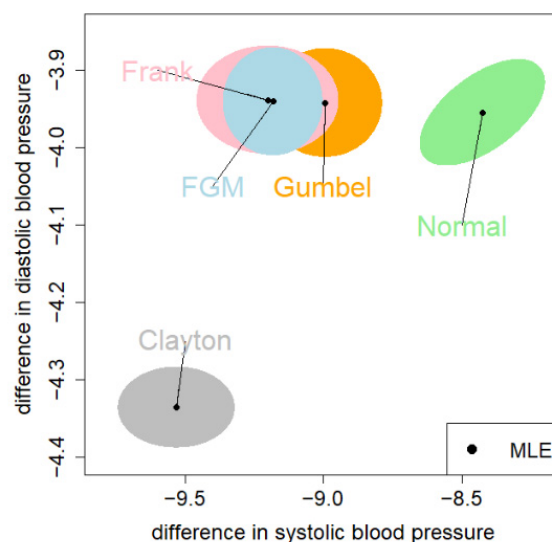

Figure 5. The MLE and the 95% CE (colored region) for the common means based on the blood pressure data. The correspondence copulas of the colored regions are: blue = FGM; gray = Clayton; orange = Gumbel; pink = Frank; green = normal.

Figure 6 depicts the 3D plot of the log-likelihood surface under the Frank copula model. The plot shows that the estimate of the common mean $\hat{\mu}_n^{\text{Frank}} = (-9.20, -3.94)$ attains the global maximum of the log-likelihood function.
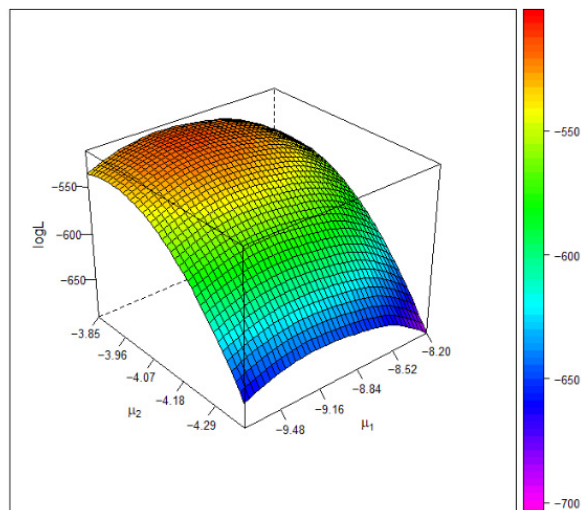

Figure 6. The 3D plot of the log-likelihood value under the Frank copula based on the blood pressure data. The maximum occurs at $\hat{\mu}_n^{\text{Frank}} = (-9.20, -3.94)$.

9. Extension to Non-Normal Models

So far, we have considered a common mean model under the marginal normality. This section explains how the proposed methods can be extended to non-normal models. For this reason, we specifically consider a common mean model under the marginal exponential distributions.

$$\bar{F}_j(y) \equiv Pr(Y_{ij} > y) = \exp(-\lambda_j y), \quad y > 0, \quad \lambda_j > 0, \quad j = 1, 2.$$

Thus, the common mean vector is $\boldsymbol{\mu} = (1/\lambda_1, 1/\lambda_2)$.

We consider the Clayton copula to specify the bivariate distribution because it has simple derivatives with respect to the copula parameter [54]. Therefore, we propose a bivariate common mean Clayton copula model with exponential margins as follows:

$$Pr(Y_{i1} > y_1, Y_{i2} > y_2) = C_{\alpha_i}^{\text{Clayton}}\left\{\bar{F}_1(y_1), \bar{F}_2(y_2)\right\} = \left\{\exp(\alpha_i \lambda_1 y_1) + \exp(\alpha_i \lambda_2 y_2) - 1\right\}^{-1/\alpha_i}, \quad (4)$$

where α_i is known for $i = 1, 2, \ldots, n$. Note that copula $C_{\alpha_i}^{\text{Clayton}}$ is a survival copula for (Y_{i1}, Y_{i2}) as the usual way to model a survival function [22]. Using similar arguments to [55], the information matrix with respect to $\boldsymbol{\lambda} = (\lambda_1, \lambda_2)$ can be decomposed as

$$I_i(\boldsymbol{\lambda}) = \begin{bmatrix} \frac{1}{\lambda_1^2} & 0 \\ 0 & \frac{1}{\lambda_2^2} \end{bmatrix} + \begin{bmatrix} \frac{2\alpha_i^2(\alpha_i+1)}{\lambda_1^2(3\alpha_i+1)} & -\frac{\alpha_i(2\alpha_i+1)}{\lambda_1\lambda_2}\phi(\alpha_i) \\ -\frac{\alpha_i(2\alpha_i+1)}{\lambda_1\lambda_2}\phi(\alpha_i) & \frac{2\alpha_i^2(\alpha_i+1)}{\lambda_2^2(3\alpha_i+1)} \end{bmatrix}, \quad (5)$$

where

$$\phi(\alpha) = \frac{1}{3\alpha+1} + \frac{1}{2(3\alpha+1)(2\alpha+1)}\left\{\Psi\left(\frac{1}{2\alpha}\right) - \Psi\left(\frac{\alpha+1}{2\alpha}\right)\right\}, \quad \Psi(\alpha) = \frac{d^2 \log \Gamma(\alpha)}{d\alpha^2}.$$

See Appendix A.3. for detailed derivations. The expression of $I_i(\boldsymbol{\lambda})$ is an extension of Theorem 3 to the exponential model. With the information matrix, the properties of the MLE and the asymptotic theory are similar to the normal models.

We conducted Monte Carlo simulations to examine the correctness of Equation (5) by comparing it with their empirical version. We set $\lambda_1 = \lambda_2 = 1$ and $\alpha_i = 1$ for all i. We generated data (Y_{i1}, Y_{i2}), $i = 1, \ldots, n$ from the model in Equation (4) and computed the empirical versions of $I_{i,11}(\boldsymbol{\lambda})$ and $I_{i,12}(\boldsymbol{\lambda})$ as

$$\frac{1}{n} \sum_{i=1}^{n} \frac{\partial^2 \log f_{i,\lambda}^{\text{Clayton}}(Y_{i1}, Y_{i2})}{\partial \lambda_1^2}, \quad \frac{1}{n} \sum_{i=1}^{n} \frac{\partial^2 \log f_{i,\lambda}^{\text{Clayton}}(Y_{i1}, Y_{i2})}{\partial \lambda_1 \partial \lambda_2}.$$

The formulas for the derivatives of the log-density are found in Equations (A1) and (A2) in Appendix A.3. Our simulations were based on 1000 repetitions with $n \in \{100, 200, 300, 400, 500\}$.

Figure 7 depicts the simulation results based on 1000 repetitions. It clearly shows that the empirical versions are scattered around the theoretical values of $I_{i,11}(\boldsymbol{\lambda}) = 2$ and $I_{i,12}(\boldsymbol{\lambda}) = -1.16$. The variability of the empirical versions vanishes as n increases. The simulation results assert the correctness of Equation (5).
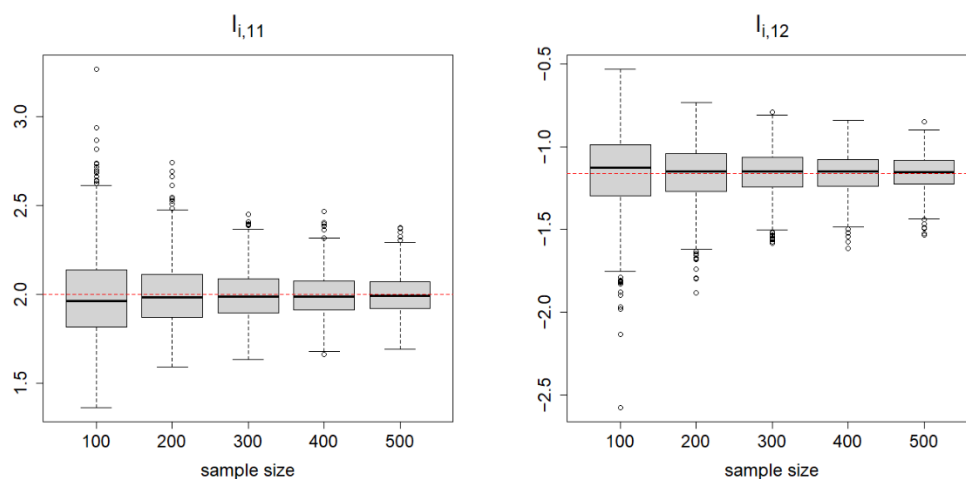

Figure 7. The boxplots summarizing the 1000 repetitions for the empirical versions of $I_{i,11}(\boldsymbol{\lambda}) = 2$ and $I_{i,12}(\boldsymbol{\lambda}) = -1.16$ (dashed lines) under the Clayton copula model in Equation (4) with parameter = 1. The sample size varies from $n = 100$ to $n = 300$.

10. Conclusions

At present, copula models are very popular in all areas of science. Bivariate meta-analyses are among those research areas that require sophisticated copula-based methods and theories. Nonetheless, there are only a few studies on copula-based bivariate meta-analysis from a methodological/theoretical perspective. This article fully develops the methodologies and theories of the copula-based bivariate meta-analysis, specifically for estimating the common mean vector. These developments will provide solid methodological/theoretical bases that are not available to date.

In this article, we emphasize the flexibility of the proposed copula models that allow for a variety of dependence structures. In the two real data examples, we employed the log-likelihood value as a criterion for model selection (Section 8). Even if the best copula is selected, it still raises the issue of goodness-of-fit, which is difficult to assess under the meta-analysis setting. The classical methods, such as Kolmogorov–Smirnov or Cramér–von Mises type statistics, cannot be directly applied to the non-identically distributed samples for which the empirical distribution function is difficult to interpret. Therefore, the development of goodness-of-fit tests is a possible research direction.

The fundamental assumption made in the proposed model is the common mean model, with known within-study correlations. The common mean assumption, although convenient for summarizing the data for a small number of studies [56], may not always hold in real meta-analyses [6]. Therefore, the extension of the proposed estimator to random

means (random-effects models) or ordered means [57,58] is an important direction for future research. To model the random effects, we need another bivariate copula. The estimation problem for these hierarchical copula-based models is beyond the scope of the paper. Nonetheless, the results presented in this paper serve as fundamental knowledge before the exploration of more advanced models.

Appendix A

Appendix A.1. Proof of Lemma 1

We first prepare a lemma:

Lemma A1. *Under the general copula model (Definition 1), if $C_{\boldsymbol{\theta}}^{[2,2]}$ exists in $(0,1)^2$, the correlation function has alternative expressions*

$$
\begin{aligned}
\rho_C(\boldsymbol{\theta}_i) &= E\left[\frac{Z_{i2}\,\varphi(Z_{i1})\,C_{\boldsymbol{\theta}_i}^{[2,1]}\{\Phi(Z_{i1}),\Phi(Z_{i2})\}}{C_{\boldsymbol{\theta}_i}^{[1,1]}\{\Phi(Z_{i1}),\Phi(Z_{i2})\}}\right] \\
&= E\left[\frac{Z_{i1}\,\varphi(Z_{i2})\,C_{\boldsymbol{\theta}_i}^{[1,2]}\{\Phi(Z_{i1}),\Phi(Z_{i2})\}}{C_{\boldsymbol{\theta}_i}^{[1,1]}\{\Phi(Z_{i1}),\Phi(Z_{i2})\}}\right] \\
&= E\left[\frac{\varphi(Z_{i1})\,\varphi(Z_{i2})\,C_{\boldsymbol{\theta}_i}^{[2,2]}\{\Phi(Z_{i1}),\Phi(Z_{i2})\}}{C_{\boldsymbol{\theta}_i}^{[1,1]}\{\Phi(Z_{i1}),\Phi(Z_{i2})\}}\right],
\end{aligned}
$$

where Z_{i1} and Z_{i2} have the joint density

$$
f_i(z_1,z_2) = \varphi(z_1)\varphi(z_2)C_{\boldsymbol{\theta}_i}^{[1,1]}\{\Phi(z_1),\Phi(z_2)\}.
$$

Proof of Lemma A1. We only prove the first identity for illustration. If $C_{\boldsymbol{\theta}}^{[2,2]}$ exists, then

$$
\begin{aligned}
\rho_C(\boldsymbol{\theta}_i) &= \int_{-\infty}^{\infty} z_2\varphi(z_2)\int_{-\infty}^{\infty} z_1\varphi(z_1)C_{\boldsymbol{\theta}_i}^{[1,1]}\{\Phi(z_1),\Phi(z_2)\}dz_1 dz_2 \\
&= \int_{-\infty}^{\infty} z_2\varphi(z_2)\int_{-\infty}^{\infty} \varphi(z_1)^2 C_{\boldsymbol{\theta}_i}^{[2,1]}\{\Phi(z_1),\Phi(z_2)\}dz_1 dz_2,
\end{aligned}
$$

where the last equality follows from Stein's identity. Thus, we obtain

$$
\begin{aligned}
\rho_C(\boldsymbol{\theta}_i) &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} z_2 \varphi(z_2) \varphi(z_1)^2 C_{\boldsymbol{\theta}_i}^{[2,1]} \{\Phi(z_1), \Phi(z_2)\} dz_1 dz_2 \\
&= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} z_2 \varphi(z_1) \frac{C_{\boldsymbol{\theta}_i}^{[2,1]} \{\Phi(z_1), \Phi(z_2)\}}{C_{\boldsymbol{\theta}_i}^{[1,1]} \{\Phi(z_1), \Phi(z_2)\}} f_i(z_1, z_2) dz_1 dz_2 \\
&= E\left[\frac{Z_{i2} \varphi(Z_{i1}) C_{\boldsymbol{\theta}_i}^{[2,1]} \{\Phi(Z_{i1}), \Phi(Z_{i2})\}}{C_{\boldsymbol{\theta}_i}^{[1,1]} \{\Phi(Z_{i1}), \Phi(Z_{i2})\}} \right].
\end{aligned}
$$

The proof completes. $\square$

Lemma A1 is a generalization of Lemma 3.2 in [3].

Now, we prove Lemma 1 for $j = 1$ and $k = 2$. If $C_{\boldsymbol{\theta}}^{[2,2]}$ exists, by straightforward calculations,

$$
\begin{aligned}
E&\left\{ \frac{\partial \log f_{i,\boldsymbol{\mu}}(\boldsymbol{Y}_i)}{\partial \mu_j} \frac{\partial \log f_{i,\boldsymbol{\mu}}(\boldsymbol{Y}_i)}{\partial \mu_k} \right\} \\
&= \frac{1}{\sigma_{i1}\sigma_{i2}} \left(E(Z_{i1}Z_{i2}) - E\left[\frac{Z_{i2}\varphi(Z_{i1})C_{\boldsymbol{\theta}_i}^{[2,1]}\{\Phi(Z_{i1}),\Phi(Z_{i2})\}}{C_{\boldsymbol{\theta}_i}^{[1,1]}\{\Phi(Z_{i1}),\Phi(Z_{i2})\}} \right] \right. \\
&\quad\quad -E\left[\frac{Z_{i1}\varphi(Z_{i2})C_{\boldsymbol{\theta}_i}^{[1,2]}\{\Phi(Z_{i1}),\Phi(Z_{i2})\}}{C_{\boldsymbol{\theta}_i}^{[1,1]}\{\Phi(Z_{i1}),\Phi(Z_{i2})\}} \right] \\
&\quad\quad \left. +E\left[\frac{\varphi(Z_{i1})\varphi(Z_{i2})C_{\boldsymbol{\theta}_i}^{[2,1]}\{\Phi(Z_{i1}),\Phi(Z_{i2})\}C_{\boldsymbol{\theta}_i}^{[1,2]}\{\Phi(Z_{i1}),\Phi(Z_{i2})\}}{C_{\boldsymbol{\theta}_i}^{[1,1]}\{\Phi(Z_{i1}),\Phi(Z_{i2})\}^2} \right] \right).
\end{aligned}
$$

On the other hand,

$$
\begin{aligned}
E&\left\{ -\frac{\partial^2 \log f_{i,\boldsymbol{\mu}}(\boldsymbol{Y}_i)}{\partial \mu_j \partial \mu_k} \right\} \\
&= \frac{1}{\sigma_{i1}\sigma_{i2}} \left(E\left[\frac{\varphi(Z_{i1})\varphi(Z_{i2})C_{\boldsymbol{\theta}_i}^{[2,1]}\{\Phi(Z_{i1}),\Phi(Z_{i2})\}C_{\boldsymbol{\theta}_i}^{[1,2]}\{\Phi(Z_{i1}),\Phi(Z_{i2})\}}{C_{\boldsymbol{\theta}_i}^{[1,1]}\{\Phi(Z_{i1}),\Phi(Z_{i2})\}^2} \right] \right. \\
&\quad\quad \left. -E\left[\frac{\varphi(Z_{i1})\varphi(Z_{i2})C_{\boldsymbol{\theta}_i}^{[2,2]}\{\Phi(Z_{i1}),\Phi(Z_{i2})\}}{C_{\boldsymbol{\theta}_i}^{[1,1]}\{\Phi(Z_{i1}),\Phi(Z_{i2})\}} \right] \right).
\end{aligned}
$$

Based on the above results, it suffices to show

$$
\begin{aligned}
E(Z_{i1}Z_{i2}) &- E\left[\frac{Z_{i2}\varphi(Z_{i1})C_{\boldsymbol{\theta}_i}^{[2,1]}\{\Phi(Z_{i1}),\Phi(Z_{i2})\}}{C_{\boldsymbol{\theta}_i}^{[1,1]}\{\Phi(Z_{i1}),\Phi(Z_{i2})\}} \right] \\
&-E\left[\frac{Z_{i1}\varphi(Z_{i2})C_{\boldsymbol{\theta}_i}^{[1,2]}\{\Phi(Z_{i1}),\Phi(Z_{i2})\}}{C_{\boldsymbol{\theta}_i}^{[1,1]}\{\Phi(Z_{i1}),\Phi(Z_{i2})\}} \right] \\
&+E\left[\frac{\varphi(Z_{i1})\varphi(Z_{i2})C_{\boldsymbol{\theta}_i}^{[2,2]}\{\Phi(Z_{i1}),\Phi(Z_{i2})\}}{C_{\boldsymbol{\theta}_i}^{[1,1]}\{\Phi(Z_{i1}),\Phi(Z_{i2})\}} \right] = 0
\end{aligned}
$$

which is asserted by Lemma A1. Hence, the proof is completed. $\square$

Appendix A.2. Derivatives for Copulas

The normal copula:

$$
C_\rho^{\text{Normal }[1,1]}(u,v) = \frac{1}{\sqrt{1-\rho^2}} \exp\left\{ -\frac{\rho^2 \Phi^{-1}(u)^2}{2(1-\rho^2)} - \frac{\rho^2 \Phi^{-1}(v)^2}{2(1-\rho^2)} + \frac{\rho \Phi^{-1}(u)\Phi^{-1}(v)}{1-\rho^2} \right\},
$$

$$C_\rho^{\text{Normal }[2,1]}(u,v) = \frac{\sqrt{2\pi}}{\sqrt{1-\rho^2}} \exp\left\{\frac{(1-2\rho^2)\Phi^{-1}(u)^2}{2(1-\rho^2)} - \frac{\rho^2\Phi^{-1}(v)^2}{2(1-\rho^2)} + \frac{\rho\Phi^{-1}(u)\Phi^{-1}(v)}{1-\rho^2}\right\}$$

$$\times\left\{\frac{\rho\Phi^{-1}(v)}{1-\rho^2} - \frac{\rho^2\Phi^{-1}(u)}{1-\rho^2}\right\},$$

$$C_\rho^{\text{Normal }[2,2]}(u,v) = \frac{2\pi}{\sqrt{1-\rho^2}} \exp\left\{\frac{(1-2\rho^2)\Phi^{-1}(u)^2}{2(1-\rho^2)} - \frac{(1-2\rho^2)\Phi^{-1}(v)^2}{2(1-\rho^2)} + \frac{\rho\Phi^{-1}(u)\Phi^{-1}(v)}{1-\rho^2}\right\}$$

$$\times\left[\left\{\frac{\rho\Phi^{-1}(u)}{1-\rho^2} - \frac{\rho^2\Phi^{-1}(v)}{1-\rho^2}\right\}\left\{\frac{\rho\Phi^{-1}(v)}{1-\rho^2} - \frac{\rho^2\Phi^{-1}(u)}{1-\rho^2}\right\} + \frac{\rho}{1-\rho^2}\right],$$

$$C_\rho^{\text{Normal }[3,1]}(u,v) = \frac{2\pi}{\sqrt{1-\rho^2}} \exp\left\{\frac{(2-3\rho^2)\Phi^{-1}(u)^2}{2(1-\rho^2)} - \frac{\rho^2\Phi^{-1}(v)^2}{2(1-\rho^2)} + \frac{\rho\Phi^{-1}(u)\Phi^{-1}(v)}{1-\rho^2}\right\}$$

$$\times\left[\left\{\frac{(1-2\rho^2)\Phi^{-1}(u)}{1-\rho^2} + \frac{\rho\Phi^{-1}(v)}{1-\rho^2}\right\}\left\{\frac{\rho\Phi^{-1}(v)}{1-\rho^2} - \frac{\rho^2\Phi^{-1}(u)}{1-\rho^2}\right\} - \frac{\rho^2}{1-\rho^2}\right],$$

$$C_\rho^{\text{Normal }[3,2]}(u,v) = \frac{(2\pi)^{3/2}}{\sqrt{1-\rho^2}} \exp\left\{\frac{(2-3\rho^2)\Phi^{-1}(u)^2}{2(1-\rho^2)} - \frac{(1-2\rho^2)\Phi^{-1}(v)^2}{2(1-\rho^2)} + \frac{\rho\Phi^{-1}(u)\Phi^{-1}(v)}{1-\rho^2}\right\}$$

$$\times\left[\left\{\frac{(1-2\rho^2)\Phi^{-1}(u)}{1-\rho^2} + \frac{\rho\Phi^{-1}(v)}{1-\rho^2}\right\}\left\{\frac{\rho\Phi^{-1}(v)}{1-\rho^2} - \frac{\rho^2\Phi^{-1}(u)}{1-\rho^2}\right\} - \frac{\rho^2}{1-\rho^2}\right.$$

$$\left. + \frac{\rho}{1-\rho^2}\left\{\frac{\rho\Phi^{-1}(v)}{1-\rho^2} + \frac{\rho^2\Phi^{-1}(u)}{1-\rho^2}\right\} + \frac{\rho}{1-\rho^2}\left\{\frac{(1-2\rho^2)\Phi^{-1}(u)}{1-\rho^2} + \frac{\rho\Phi^{-1}(v)}{1-\rho^2}\right\}\right],$$

$$C_\rho^{\text{Normal }[4,1]}(u,v) = \frac{(2\pi)^{3/2}}{\sqrt{1-\rho^2}} \exp\left\{\frac{(3-4\rho^2)\Phi^{-1}(u)^2}{2(1-\rho^2)} - \frac{\rho^2\Phi^{-1}(v)^2}{2(1-\rho^2)} + \frac{\rho\Phi^{-1}(u)\Phi^{-1}(v)}{1-\rho^2}\right\}$$

$$\times\left[\left\{\frac{(1-2\rho^2)\Phi^{-1}(u)}{1-\rho^2} + \frac{\rho\Phi^{-1}(v)}{1-\rho^2}\right\}\left\{\frac{\rho\Phi^{-1}(v)}{1-\rho^2} - \frac{\rho^2\Phi^{-1}(u)}{1-\rho^2}\right\} - \frac{\rho^2}{1-\rho^2}\right.$$

$$\left. + \frac{1-2\rho^2}{1-\rho^2}\left\{\frac{\rho\Phi^{-1}(v)}{1-\rho^2} + \frac{\rho^2\Phi^{-1}(u)}{1-\rho^2}\right\} - \frac{\rho^2}{1-\rho^2}\left\{\frac{(1-2\rho^2)\Phi^{-1}(u)}{1-\rho^2} + \frac{\rho\Phi^{-1}(v)}{1-\rho^2}\right\}\right].$$

The FGM copula:

$$C_\theta^{\text{FGM }[1,1]}(u,v) = 1 + \theta(1-2u)(1-2v), \quad C_\theta^{\text{FGM }[2,1]}(u,v) = -2\theta(1-2v),$$

$$C_\theta^{\text{FGM }[2,2]}(u,v) = 4\theta, \quad C_\theta^{\text{FGM }[3,1]}(u,v) = C_\theta^{\text{FGM }[3,2]}(u,v) = C_\theta^{\text{FGM }[4,1]}(u,v) = 0.$$

The Clayton copula:

$$C_\alpha^{\text{Clayton }[j,k]}(u,v) = (\alpha+1)u^{-\alpha-j}v^{-\alpha-k}(u^{-\alpha} + v^{-\alpha} - 1)^{-1/\alpha-(j+k)}\omega_{j,k}(u,v),$$

$$\text{where } (j,k) \in \{(1,1),(2,1),(2,2),(3,1),(3,2),(4,1)\},$$

$$\omega_{1,1}(u,v) = 1, \quad \omega_{2,1}(u,v) = \alpha u^{-\alpha} + (\alpha+1)(1-v^{-\alpha}),$$

$$\omega_{2,2}(u,v) = -\alpha(\alpha+1)u^{-2\alpha} - (\alpha+1)u^{-\alpha} + \left(4\alpha^2 + 3\alpha + 1\right)u^{-\alpha}v^{-\alpha} - \alpha(\alpha+1)v^{-2\alpha} - (\alpha+1)v^{-\alpha} + (\alpha+1)^2.$$

$$\omega_{3,1}(u,v) = \alpha(\alpha-1)u^{-2\alpha} + (4\alpha-1)(\alpha+1)(1-v^{-\alpha})u^{-\alpha} + (\alpha+1)(\alpha+2)(1-v^{-\alpha})^2,$$

$$\omega_{3,2}(u,v) = -\alpha(\alpha-1)(\alpha+1)u^{-3\alpha} - (\alpha+1)\left(3\alpha^2+4\alpha-1\right)u^{-2\alpha} + 3(\alpha-1)(\alpha+1)^2u^{-\alpha}$$

$$+ \alpha(\alpha+1)(\alpha+2)v^{-3\alpha} - (\alpha-1)(\alpha+1)(\alpha+2)v^{-2\alpha} - (\alpha+1)(\alpha+2)^2v^{-\alpha}$$

$$+ \left(11\alpha^3+7\alpha^2+\alpha-1\right)u^{-2\alpha}v^{-\alpha} - (\alpha+1)\left(11\alpha^2+5\alpha+2\right)u^{-\alpha}v^{-2\alpha}$$

$$+ (\alpha+1)\left(8\alpha^2+5\alpha+5\right)u^{-\alpha}v^{-\alpha} + (\alpha+1)^2(\alpha+2).$$

$$\omega_{4,1}(u,v) = \alpha(\alpha-1)(\alpha-2)u^{-3\alpha} + (\alpha-1)(\alpha+1)(11\alpha-2)\left(1-v^{-\alpha}\right)u^{-2\alpha}$$

$$+ (\alpha+1)\left(11\alpha^2+14\alpha-7\right)\left(1-v^{-\alpha}\right)^2u^{-\alpha} + (\alpha+1)(\alpha+2)(\alpha+3)\left(1-v^{-\alpha}\right)^3.$$

Appendix A.3. The Information Matrix under the Clayton Copula with Exponential Margins

The Clayton copula model with exponential margins is given as

$$Pr(Y_{i1}>y_1, Y_{i2}>y_2) = C_{\alpha_i}^{\text{Clayton}}\left\{\bar{F}_1(y_1), \bar{F}_2(y_2)\right\} = \left\{\exp(\alpha_i\lambda_1 y_1) + \exp(\alpha_i\lambda_2 y_2) - 1\right\}^{-1/\alpha_i}.$$

Then, the joint density is

$$f_{i,\lambda}^{\text{Clayton}}(y_1, y_2) = \frac{\partial^2}{\partial y_1 \partial y_2}Pr(Y_{i1}>y_1, Y_{i2}>y_2) = \frac{(\alpha_i+1)\lambda_1\lambda_2 e^{\alpha_i\lambda_1 y_1 + \alpha_i\lambda_2 y_2}}{\left(e^{\alpha_i\lambda_1 y_1} + e^{\alpha_i\lambda_2 y_2} - 1\right)^{1/\alpha_i+2}},$$

where $\lambda = (\lambda_1, \lambda_2)$. The log-density is

$$\log f_{i,\lambda}^{\text{Clayton}}(y_1, y_2) = \log(\alpha_i+1) + \log\lambda_1 + \log\lambda_2 + \alpha_i\lambda_1 y_1 + \alpha_i\lambda_2 y_2 - \left(\frac{1}{\alpha_i}+2\right)\log\left(e^{\alpha_i\lambda_1 y_1} + e^{\alpha_i\lambda_2 y_2} - 1\right).$$

The first-order partial derivative of $\log f_{i,\lambda}^{\text{Clayton}}(y_1, y_2)$ with respect to λ_j is

$$\frac{\partial \log f_{i,\lambda}^{\text{Clayton}}(y_1, y_2)}{\partial \lambda_j} = \frac{1}{\lambda_j} + \alpha_i y_j - (2\alpha_i+1)\frac{y_j e^{\alpha_i\lambda_j y_j}}{e^{\alpha_i\lambda_1 y_1} + e^{\alpha_i\lambda_2 y_2} - 1}, \quad j=1,2.$$

The second-order partial derivatives of $\log f_{\alpha_i}^{\text{Clayton}}(y_1, y_2)$ are

$$\frac{\partial^2 \log f_{i,\lambda}^{\text{Clayton}}(y_1, y_2)}{\partial \lambda_j^2} = -\frac{1}{\lambda_j^2} - \alpha_i(2\alpha_i+1)\left\{\frac{y_j^2 e^{\alpha_i\lambda_j y_j}}{e^{\alpha_i\lambda_1 y_1} + e^{\alpha_i\lambda_2 y_2} - 1} - \frac{y_j^2 e^{2\alpha_i\lambda_j y_j}}{\left(e^{\alpha_i\lambda_1 y_1} + e^{\alpha_i\lambda_2 y_2} - 1\right)^2}\right\}, \quad j=1,2, \tag{A1}$$

$$\frac{\partial^2 \log f_{i,\lambda}^{\text{Clayton}}(y_1, y_2)}{\partial \lambda_1 \partial \lambda_2} = \alpha_i(2\alpha_i+1)\frac{y_1 y_2 e^{\alpha_i\lambda_1 y_1 + \alpha_i\lambda_2 y_2}}{\left(e^{\alpha_i\lambda_1 y_1} + e^{\alpha_i\lambda_2 y_2} - 1\right)^2}. \tag{A2}$$

Then, the Fisher information matrix is

$$I_{i,jk}^{\text{Clayton}}(\lambda) = E\left\{-\frac{\partial^2 \log f_{i,\lambda}^{\text{Clayton}}(Y_i)}{\partial \lambda_j \partial \lambda_k}\right\}, \quad j,k=1,2.$$

For $j=k=1$, we have

$$I_{i,11}^{\text{Clayton}}(\lambda) = E\left\{-\frac{\partial^2 \log f_{i,\lambda}^{\text{Clayton}}(Y_i)}{\partial \lambda_1^2}\right\}$$

$$= \frac{1}{\lambda_1^2} + \frac{\alpha_i(2\alpha_i + 1)}{\lambda_1^2}\left[E\left(\frac{\lambda_1^2 Y_{i1}^2 e^{\alpha_i\lambda_1 Y_{i1}}}{e^{\alpha_i\lambda_1 Y_{i1}} + e^{\alpha_i\lambda_2 Y_{i2}} - 1}\right) - E\left\{\frac{\lambda_1^2 Y_{i1}^2 e^{2\alpha_i\lambda_1 Y_{i1}}}{\left(e^{\alpha_i\lambda_1 Y_{i1}} + e^{\alpha_i\lambda_2 Y_{i2}} - 1\right)^2}\right\}\right].$$

To compute the first expectation, we consider the change in variable $\lambda_j y_j = x_j$, $j = 1, 2$.
Then,

$$E\left(\frac{\lambda_1^2 Y_{i1}^2 e^{\alpha_i\lambda_1 Y_{i1}}}{e^{\alpha_i\lambda_1 Y_{i1}} + e^{\alpha_i\lambda_2 Y_{i2}} - 1}\right) = (\alpha_i + 1)\int_0^\infty \int_0^\infty \frac{\lambda_1^3 y_1^2 \lambda_2 e^{2\alpha_i\lambda_1 y_1 + \alpha_i\lambda_2 y_2}}{\left(e^{\alpha_i y_1} + e^{\alpha_i y_2} - 1\right)^{1/\alpha_i + 3}}dy_1 dy_2$$

$$= (\alpha_i + 1)\int_0^\infty \int_0^\infty \frac{x_1^2 e^{2\alpha_i x_1 + \alpha_i x_2}}{\left(e^{\alpha_i x_1} + e^{\alpha_i x_2} - 1\right)^{1/\alpha_i + 3}}dx_1 dx_2$$

$$= (\alpha_i + 1)\int_0^\infty x_1^2 e^{2\alpha_i x_1}\int_0^\infty \frac{e^{\alpha_i x_2}}{\left(e^{\alpha_i x_1} + e^{\alpha_i x_2} - 1\right)^{1/\alpha_i + 3}}dx_1 dx_2.$$

For the inner integral,

$$\int_0^\infty \frac{e^{\alpha_i x_2}}{\left(e^{\alpha_i x_1} + e^{\alpha_i x_2} - 1\right)^{1/\alpha_i + 3}}dx_2 = -\frac{1}{2\alpha_i + 1}\left\{\frac{1}{\left(e^{\alpha_i x_1} + e^{\alpha_i x_2} - 1\right)^{1/\alpha_i + 2}}\bigg|_0^\infty\right\} = \frac{1}{2\alpha_i + 1}e^{-(2\alpha_i + 1)x_1}.$$

It follows that

$$\int_0^\infty x_1^2 e^{2\alpha_i x_1}\left\{\frac{1}{2\alpha_i + 1}e^{-(2\alpha_i + 1)x_1}\right\}dx_1 = \frac{1}{2\alpha_i + 1}\int_0^\infty x_1^2 e^{-x_1}dx_1 = \frac{2}{2\alpha_i + 1}.$$

Thus, we obtain

$$E\left(\frac{\lambda_1^2 Y_{i1}^2 e^{\alpha_i\lambda_1 Y_{i1}}}{e^{\alpha_i\lambda_1 Y_{i1}} + e^{\alpha_i\lambda_2 Y_{i2}} - 1}\right) = \frac{2(\alpha_i + 1)}{2\alpha_i + 1}.$$

The second expectation is computed as

$$E\left\{\frac{\lambda_1^2 Y_{i1}^2 e^{2\alpha_i\lambda_1 Y_{i1}}}{\left(e^{\alpha_i\lambda_1 Y_{i1}} + e^{\alpha_i\lambda_2 Y_{i2}} - 1\right)^2}\right\} = (\alpha_i + 1)\int_0^\infty \int_0^\infty \frac{\lambda_1^3 y_1^2 \lambda_2 e^{3\alpha_i\lambda_1 y_1 + \alpha_i\lambda_2 y_2}}{\left(e^{\alpha_i y_1} + e^{\alpha_i y_2} - 1\right)^{1/\alpha_i + 4}}dy_1 dy_2$$

$$= (\alpha_i + 1)\int_0^\infty \int_0^\infty \frac{x_1^2 e^{3\alpha_i x_1 + \alpha_i x_2}}{\left(e^{\alpha_i x_1} + e^{\alpha_i x_2} - 1\right)^{1/\alpha_i + 4}}dx_1 dx_2$$

$$= (\alpha_i + 1)\int_0^\infty x_1^2 e^{3\alpha_i x_1}\int_0^\infty \frac{e^{\alpha_i x_2}}{\left(e^{\alpha_i x_1} + e^{\alpha_i x_2} - 1\right)^{1/\alpha_i + 4}}dx_1 dx_2.$$

For the inner integral,

$$\int_0^\infty \frac{e^{\alpha_i x_2}}{\left(e^{\alpha_i x_1} + e^{\alpha_i x_2} - 1\right)^{1/\alpha_i + 4}}dx_2 = -\frac{1}{3\alpha_i + 1}\left\{\frac{1}{\left(e^{\alpha_i x_1} + e^{\alpha_i x_2} - 1\right)^{1/\alpha_i + 3}}\bigg|_0^\infty\right\} = \frac{1}{3\alpha_i + 1}e^{-(3\alpha_i + 1)x_1}.$$

It follows that

$$\int_0^\infty x_1^2 e^{3\alpha_i x_1}\left\{\frac{1}{3\alpha_i + 1}e^{-(3\alpha_i + 1)x_1}\right\}dx_1 = \frac{1}{3\alpha_i + 1}\int_0^\infty x_1^2 e^{-x_1}dx_1 = \frac{2}{3\alpha_i + 1}.$$

Thus, we obtain

$$E\left\{\frac{\lambda_1^2 Y_{i1}^2 e^{2\alpha_i\lambda_1 Y_{i1}}}{\left(e^{\alpha_i\lambda_1 Y_{i1}} + e^{\alpha_i\lambda_2 Y_{i2}} - 1\right)^2}\right\} = \frac{2(\alpha_i + 1)}{3\alpha_i + 1}.$$

Combining the above results, we have

$$
I_{i,11}^{\text{Clayton}}(\boldsymbol{\lambda}) = \frac{1}{\lambda_1^2} + \frac{1}{\lambda_1^2} \frac{2\alpha_i^2(\alpha_i + 1)}{3\alpha_i + 1}.
$$

In a similar fashion, for $j = k = 2$, we also have

$$
I_{i,22}^{\text{Clayton}}(\boldsymbol{\lambda}) = \frac{1}{\lambda_2^2} + \frac{1}{\lambda_2^2} \frac{2\alpha_i^2(\alpha_i + 1)}{3\alpha_i + 1}.
$$

For $j = 1, k = 2$, we have

$$
I_{i,12}^{\text{Clayton}}(\boldsymbol{\lambda}) = E\left\{ -\frac{\partial^2 \log f_{i,\boldsymbol{\lambda}}^{\text{Clayton}}(\boldsymbol{Y}_i)}{\partial \lambda_1 \partial \lambda_2} \right\} = -\left(\frac{1}{\alpha_i} + 2 \right) E\left\{ \frac{\lambda_1 \lambda_2 Y_{i1} Y_{i2} e^{\alpha_i \lambda_1 Y_{i1} + \alpha_i \lambda_2 Y_{i2}}}{\left(e^{\alpha_i \lambda_1 Y_{i1}} + e^{\alpha_i \lambda_2 Y_{i2}} - 1 \right)^2} \right\}
$$

$$
= -\frac{\alpha_i(2\alpha_i + 1)}{\lambda_1 \lambda_2} E\left\{ \frac{\lambda_1 \lambda_2 Y_{i1} Y_{i2} e^{\alpha_i \lambda_1 Y_{i1} + \alpha_i \lambda_2 Y_{i2}}}{\left(e^{\alpha_i \lambda_1 Y_{i1}} + e^{\alpha_i \lambda_2 Y_{i2}} - 1 \right)^2} \right\}.
$$

We consider

$$
E\left\{ \frac{\lambda_1 \lambda_2 Y_{i1} Y_{i2} e^{\alpha_i \lambda_1 Y_{i1} + \alpha_i \lambda_2 Y_{i2}}}{\left(e^{\alpha_i \lambda_1 Y_{i1}} + e^{\alpha_i \lambda_2 Y_{i2}} - 1 \right)^2} \right\} = (\alpha_i + 1) \int_0^\infty \int_0^\infty \frac{\lambda_1^2 \lambda_2^2 y_1 y_2 e^{2\alpha_i \lambda_1 y_1 + 2\alpha_i \lambda_2 y_2}}{\left(e^{\alpha_i y_1} + e^{\alpha_i y_2} - 1 \right)^{1/\alpha_i + 4}} dy_1 dy_2
$$

$$
= (\alpha_i + 1) \int_0^\infty \int_0^\infty \frac{x_1 x_2 e^{2\alpha_i x_1 + 2\alpha_i x_2}}{\left(e^{\alpha_i x_1} + e^{\alpha_i x_2} - 1 \right)^{1/\alpha_i + 4}} dx_1 dx_2
$$

$$
= (\alpha_i + 1) \int_0^\infty x_1 e^{2\alpha_i x_1} \int_0^\infty \frac{x_2 e^{2\alpha_i x_2}}{\left(e^{\alpha_i x_1} + e^{\alpha_i x_2} - 1 \right)^{1/\alpha_i + 4}} dx_1 dx_2.
$$

For the inner integral,

$$
\int_0^\infty \frac{x_2 e^{2\alpha_i x_2}}{\left(e^{\alpha_i x_1} + e^{\alpha_i x_2} - 1 \right)^{1/\alpha_i + 4}} dx_1
$$

$$
= -\frac{1}{3\alpha_i + 1} \int_0^\infty x_2 e^{\alpha_i x_2} d\left\{ \frac{1}{\left(e^{\alpha_i x_1} + e^{\alpha_i x_2} - 1 \right)^{1/\alpha_i + 3}} \right\}
$$

$$
= -\frac{1}{3\alpha_i + 1} \left\{ \left. \frac{x_2 e^{\alpha_i x_2}}{\left(e^{\alpha_i x_1} + e^{\alpha_i x_2} - 1 \right)^{1/\alpha_i + 3}} \right|_0^\infty - \int_0^\infty \frac{1}{\left(e^{\alpha_i x_1} + e^{\alpha_i x_2} - 1 \right)^{1/\alpha_i + 3}} d(x_2 e^{\alpha_i x_2}) \right\}
$$

$$
= \frac{1}{3\alpha_i + 1} \int_0^\infty \frac{e^{\alpha_i x_2} + \alpha_i x_2 e^{\alpha_i x_2}}{\left(e^{\alpha_i x_1} + e^{\alpha_i x_2} - 1 \right)^{1/\alpha_i + 3}} dx_2
$$

$$
= \frac{1}{3\alpha_i + 1} \int_0^\infty \frac{e^{\alpha_i x_2}}{\left(e^{\alpha_i x_1} + e^{\alpha_i x_2} - 1 \right)^{1/\alpha_i + 3}} dx_2 + \frac{\alpha_i}{3\alpha_i + 1} \int_0^\infty \frac{x_2 e^{\alpha_i x_2}}{\left(e^{\alpha_i x_1} + e^{\alpha_i x_2} - 1 \right)^{1/\alpha_i + 3}} dx_2,
$$

where the second equality follows from integration by parts. For the first integral,

$$
\frac{1}{3\alpha_i + 1} \int_0^\infty \frac{e^{\alpha_i x_2}}{\left(e^{\alpha_i x_1} + e^{\alpha_i x_2} - 1 \right)^{1/\alpha_i + 3}} dx_2
$$

$$
= -\frac{1}{(3\alpha_i + 1)(2\alpha_i + 1)} \left(\left. \frac{1}{\left(e^{\alpha_i x_1} + e^{\alpha_i x_2} - 1 \right)^{1/\alpha_i + 2}} \right|_0^\infty \right)
$$

$$
= \frac{1}{(3\alpha_i + 1)(2\alpha_i + 1)} e^{-2(\alpha_i + 1)x_1}.
$$

The second integral is

$$\frac{\alpha_i}{3\alpha_i+1}\int_0^\infty \frac{x_2 e^{\alpha_i x_2}}{\left(e^{\alpha_i x_1}+e^{\alpha_i x_2}-1\right)^{1/\alpha_i+3}}dx_2$$

$$=-\frac{\alpha_i}{(3\alpha_i+1)(2\alpha_i+1)}\int_0^\infty x_2 d\left(\frac{1}{\left(e^{\alpha_i x_1}+e^{\alpha_i x_2}-1\right)^{1/\alpha_i+2}}\right)$$

$$=-\frac{\alpha_i}{(3\alpha_i+1)(2\alpha_i+1)}\left\{\left.\frac{x_2}{\left(e^{\alpha_i x_1}+e^{\alpha_i x_2}-1\right)^{1/\alpha_i+2}}\right|_0^\infty -\int_0^\infty \frac{1}{\left(e^{\alpha_i x_1}+e^{\alpha_i x_2}-1\right)^{1/\alpha_i+2}}dx_2\right\}$$

$$=\frac{\alpha_i}{(3\alpha_i+1)(2\alpha_i+1)}\int_0^\infty \frac{1}{\left(e^{\alpha_i x_1}+e^{\alpha_i x_2}-1\right)^{1/\alpha_i+2}}dx_2,$$

where the second equality follows from integration by parts. Now, the expectation becomes

$$E\left\{\frac{\lambda_1\lambda_2 Y_{i1}Y_{i2}e^{\alpha_i\lambda_1 Y_{i1}+\alpha_i\lambda_2 Y_{i2}}}{\left(e^{\alpha_i\lambda_1 Y_{i1}}+e^{\alpha_i\lambda_2 Y_{i2}}-1\right)^2}\right\}$$

$$=\frac{\alpha_i+1}{(3\alpha_i+1)(2\alpha_i+1)}\int_0^\infty x_1 e^{-x_1}dx_1+\frac{\alpha_i(\alpha_i+1)}{(3\alpha_i+1)(2\alpha_i+1)}\int_0^\infty\int_0^\infty \frac{x_1 e^{2\alpha_i x_2}}{\left(e^{\alpha_i x_1}+e^{\alpha_i x_2}-1\right)^{1/\alpha_i+2}}dx_1 dx_2,$$

$$=\frac{\alpha_i+1}{(3\alpha_i+1)(2\alpha_i+1)}+\frac{\alpha_i(\alpha_i+1)}{(3\alpha_i+1)(2\alpha_i+1)}\int_0^\infty\int_0^\infty \frac{x_1 e^{2\alpha_i x_2}}{\left(e^{\alpha_i x_1}+e^{\alpha_i x_2}-1\right)^{1/\alpha_i+2}}dx_1 dx_2.$$

For the integral in the above expression, we consider its inner integral,

$$\int_0^\infty \frac{x_1 e^{2\alpha_i x_2}}{\left(e^{\alpha_i x_1}+e^{\alpha_i x_2}-1\right)^{1/\alpha_i+2}}dx_1$$

$$=-\frac{1}{\alpha_i+1}\int_0^\infty x_1 e^{\alpha_i x_1}d\left\{\frac{1}{\left(e^{\alpha_i x_1}+e^{\alpha_i x_2}-1\right)^{1/\alpha_i+1}}\right\}$$

$$=-\frac{1}{\alpha_i+1}\left\{\left.\frac{x_1 e^{\alpha_i x_1}}{\left(e^{\alpha_i x_1}+e^{\alpha_i x_2}-1\right)^{1/\alpha_i+1}}\right|_0^\infty -\int_0^\infty \frac{1}{\left(e^{\alpha_i x_1}+e^{\alpha_i x_2}-1\right)^{1/\alpha_i+1}}d\left(x_1 e^{\alpha_i x_1}\right)\right\}$$

$$=\frac{1}{\alpha_i+1}\int_0^\infty \frac{e^{\alpha_i x_2}+\alpha_i x_1 e^{\alpha_i x_1}}{\left(e^{\alpha_i x_1}+e^{\alpha_i x_2}-1\right)^{1/\alpha_i+1}}dx_1$$

$$=\frac{1}{\alpha_i+1}\int_0^\infty \frac{e^{\alpha_i x_2}}{\left(e^{\alpha_i x_1}+e^{\alpha_i x_2}-1\right)^{1/\alpha_i+1}}dx_1+\frac{\alpha_i}{\alpha_i+1}\int_0^\infty \frac{x_1 e^{\alpha_i x_1}}{\left(e^{\alpha_i x_1}+e^{\alpha_i x_2}-1\right)^{1/\alpha_i+1}}dx_1,$$

where the second last equality follows from integration by parts. We compute the above two integrals separately. We have

$$\frac{1}{\alpha_i+1}\int_0^\infty \frac{e^{\alpha_i x_2}}{\left(e^{\alpha_i x_1}+e^{\alpha_i x_2}-1\right)^{1/\alpha_i+1}}dx_1=-\frac{1}{\alpha_i+1}\left\{\left.\frac{1}{\left(e^{\alpha_i x_1}+e^{\alpha_i x_2}-1\right)^{1/\alpha_i}}\right|_0^\infty\right\}=\frac{1}{\alpha_i+1}e^{-x_2}.$$

On the other hand, we have

$$\frac{\alpha_i}{\alpha_i+1}\int_0^\infty \frac{x_1 e^{\alpha_i x_1}}{\left(e^{\alpha_i x_1}+e^{\alpha_i x_2}-1\right)^{1/\alpha_i+1}}dx_1$$

$$=-\frac{\alpha_i}{\alpha_i+1}\int_0^\infty x_1 d\left\{\frac{1}{\left(e^{\alpha_i x_1}+e^{\alpha_i x_2}-1\right)^{1/\alpha_i}}\right\}$$

$$= -\frac{\alpha_i}{\alpha_i + 1}\left\{ \frac{x_1}{(e^{\alpha_i x_1} + e^{\alpha_i x_2} - 1)^{1/\alpha_i}}\bigg|_0^\infty - \int_0^\infty \frac{1}{(e^{\alpha_i x_1} + e^{\alpha_i x_2} - 1)^{1/\alpha_i}}dx_1 \right\}$$

$$= \frac{\alpha_i}{\alpha_i + 1}\int_0^\infty \frac{1}{(e^{\alpha_i x_1} + e^{\alpha_i x_2} - 1)^{1/\alpha_i}}dx_1.$$

Then,

$$\int_0^\infty \left\{ \frac{1}{\alpha_i + 1}e^{-x_2} + \frac{\alpha_i}{\alpha_i + 1}\int_0^\infty \frac{1}{(e^{\alpha_i x_1} + e^{\alpha_i x_2} - 1)^{1/\alpha_i}}dx_1 \right\}dx_2 = \frac{1}{\alpha_i + 1} + \frac{\alpha_i}{\alpha_i + 1}\int_0^\infty \int_0^\infty \frac{1}{(e^{\alpha_i x_1} + e^{\alpha_i x_2} - 1)^{1/\alpha_i}}dx_1 dx_2.$$

Combine all the results, one has

$$E\left\{ \frac{\lambda_1 \lambda_2 Y_{i1} Y_{i2} e^{\alpha_i \lambda_1 Y_{i1} + \alpha_i \lambda_2 Y_{i2}}}{(e^{\alpha_i \lambda_1 Y_{i1}} + e^{\alpha_i \lambda_2 Y_{i2}} - 1)^2} \right\}$$

$$= \frac{\alpha_i + 1}{(3\alpha_i + 1)(2\alpha_i + 1)} + \frac{\alpha_i}{(3\alpha_i + 1)(2\alpha_i + 1)} + \frac{\alpha_i^2}{(3\alpha_i + 1)(2\alpha_i + 1)}\int_0^\infty \int_0^\infty \frac{1}{(e^{\alpha_i x_1} + e^{\alpha_i x_2} - 1)^{1/\alpha_i}}dx_1 dx_2$$

$$= \frac{1}{3\alpha_i + 1} + \frac{\alpha_i^2}{(3\alpha_i + 1)(2\alpha_i + 1)}\int_0^\infty \int_0^\infty \frac{1}{(e^{\alpha_i x_1} + e^{\alpha_i x_2} - 1)^{1/\alpha_i}}dx_1 dx_2.$$

Let $\alpha_i x_1 = s$ and $\alpha_i x_2 = t$, according to [52], we have

$$\int_0^\infty \int_0^\infty \frac{\alpha_i^2}{(e^{\alpha_i x_1} + e^{\alpha_i x_2} - 1)^{1/\alpha_i}}dx_1 dx_2 = \int_0^\infty \int_0^\infty \frac{1}{(e^s + e^t - 1)^{1/\alpha_i}}ds\,dt = \frac{1}{2}\left\{ \Psi\left(\frac{1}{2\alpha_i}\right) - \Psi\left(\frac{\alpha_i + 1}{2\alpha_i}\right) \right\}.$$

Hence, we obtain

$$E\left\{ \frac{\lambda_1 \lambda_2 Y_{i1} Y_{i2} e^{\alpha_i \lambda_1 Y_{i1} + \alpha_i \lambda_2 Y_{i2}}}{(e^{\alpha_i \lambda_1 Y_{i1}} + e^{\alpha_i \lambda_2 Y_{i2}} - 1)^2} \right\} = \frac{1}{3\alpha_i + 1} + \frac{1}{2(3\alpha_i + 1)(2\alpha_i + 1)}\left\{ \Psi\left(\frac{1}{2\alpha_i}\right) - \Psi\left(\frac{\alpha_i + 1}{2\alpha_i}\right) \right\}.$$

Finally, combining the above results, we have

$$I_{i,12}^{\text{Clayton}}(\boldsymbol{\lambda}) = -\frac{\alpha_i(2\alpha_i + 1)}{\lambda_1 \lambda_2}\phi(\alpha_i),$$

where

$$\phi(\alpha) = \frac{1}{3\alpha_i + 1} + \frac{1}{2(3\alpha + 1)(2\alpha + 1)}\left\{ \Psi\left(\frac{1}{2\alpha}\right) - \Psi\left(\frac{\alpha + 1}{2\alpha}\right) \right\}.$$

References

1. Gleser, L.J.; Olkin, L. Stochastically dependent effect sizes. In *The Handbook of Research Synthesis*; Russel Sage Foundation: New York, NY, USA, 1994.
2. Riley, R.D. Multivariate meta-analysis: The effect of ignoring within-study correlation. *J. R. Stat. Soc. Ser. A Stat. Soc.* **2009**, *172*, 789–811. [CrossRef]
3. Shih, J.-H.; Konno, Y.; Chang, Y.-T.; Emura, T. Estimation of a common mean vector in bivariate meta-analysis under the FGM copula. *Statistics* **2019**, *53*, 673–695. [CrossRef]
4. Nissen, S.E.; Wolski, K. Effect of Rosiglitazone on the Risk of Myocardial Infarction and Death from Cardiovascular Causes. *N. Engl. J. Med.* **2007**, *356*, 2457–2471. [CrossRef] [PubMed]
5. Yamaguchi, Y.; Maruo, K. Bivariate beta-binomial model using Gaussian copula for bivariate meta-analysis of two binary outcomes with low incidence. *Jpn. J. Stat. Data Sci.* **2019**, *2*, 347–373. [CrossRef]
6. Mavridis, D.; Salanti, G. A practical introduction to multivariate meta-analysis. *Stat. Methods Med. Res.* **2011**, *22*, 133–158. [CrossRef]
7. Copas, J.B.; Jackson, D.; White, I.; Riley, R.D. The role of secondary outcomes in multivariate meta-analysis. *J. R. Stat. Soc. Ser. C Appl. Stat.* **2018**, *67*, 1177–1205. [CrossRef]

8. Burzykowski, T.; Molenberghs, G.; Buyse, M.; Geys, H.; Renard, D. Validation of surrogate end points in multiple randomized clinical trials with failure time end points. *J. R. Stat. Soc. Ser. C Appl. Stat.* **2001**, *50*, 405–422. [CrossRef]

9. Rotolo, F.; Paoletti, X.; Michiels, S. surrosurv: An R package for the evaluation of failure time surrogate endpoints in individual patient data meta-analysis of randomized clinical trials. *Comput. Methods Programs Biomed.* **2018**, *155*, 189–198. [CrossRef]

10. Rotolo, F.; Paoletti, X.; Burzykowski, T.; Buyse, M.; Michiels, S. A Poisson approach to the validation of failure time surrogate endpoints in individual patient data meta-analyses. *Stat. Methods Med. Res.* **2019**, *28*, 170–183. [CrossRef]

11. Emura, T.; Sofeu, C.L.; Rondeau, V. Conditional copula models for correlated survival endpoints: Individual patient data meta-analysis of randomized controlled trials. *Stat. Methods Med. Res.* **2021**, *30*, 2634–2650. [CrossRef]

12. Kuss, O.; Hoyer, A.; Solms, A. Meta-analysis for diagnostic accuracy studies: A new statistical model using beta-binomial dis-tributions and bivariate copulas. *Stat. Med.* **2014**, *33*, 17–30. [CrossRef]

13. Nikoloulopoulos, A.K. A vine copula mixed effect model for trivariate meta-analysis of diagnostic test accuracy studies ac-counting for disease prevalence. *Stat. Methods Med. Res.* **2017**, *26*, 2270–2286. [CrossRef]

14. Nikoloulopoulos, A.K. A D-vine copula mixed model for joint meta-analysis and comparison of diagnostic tests. *Stat. Methods Med. Res.* **2018**, *28*, 3286–3300. [CrossRef]

15. Nikoloulopoulos, A.K. A multinomial quadrivariate D-vine copula mixed model for meta-analysis of diagnostic studies in the presence of non-evaluable subjects. *Stat. Methods Med. Res.* **2020**, *29*, 2988–3005. [CrossRef]

16. Takeuchi, T.T. Constructing a bivariate distribution function with given marginals and correlation: Application to the galaxy luminosity function. *Mon. Not. R. Astron. Soc.* **2010**, *406*, 1830–1840. [CrossRef]

17. Ota, S.; Kimura, M. Effective estimation algorithm for parameters of multivariate Farlie–Gumbel–Morgenstern copula. *Jpn. J. Stat. Data Sci.* **2021**, *4*, 1049–1078. [CrossRef]

18. Ghosh, S.; Sheppard, L.W.; Holder, M.T.; Loecke, T.D.; Reid, P.C.; Bever, J.D.; Reuman, D.C. Copulas and their potential for ecology. *Adv. Ecol. Res.* **2020**, *62*, 409–468. [CrossRef]

19. Alidoost, F.; Stein, A.; Su, Z. The use of bivariate copulas for bias correction of reanalysis air temperature data. *PLoS ONE* **2019**, *14*, e0216059. [CrossRef]

20. Bhatti, M.I.; Do, H.Q. Recent development in copula and its applications to the energy, forestry and environmental sciences. *Int. J. Hydrog. Energy* **2019**, *44*, 19453–19473. [CrossRef]

21. Emura, T.; Chen, Y.H. *Analysis of Survival Data with Dependent Censoring: Copula-Based Approaches*; JSS Research Series in Statistics; Springer: Singapore, 2018.

22. Emura, T.; Matsui, S.; Rondeau, V. *Survival Analysis with Correlated Endpoints, Joint Frailty-Copula Models*; JSS Research Series in Statistics; Springer: Singapore, 2019.

23. Emura, T.; Shih, J.-H.; Ha, I.D.; Wilke, R.A. Comparison of the marginal hazard model and the sub-distribution hazard model for competing risks under an assumed copula. *Stat. Methods Med. Res.* **2020**, *29*, 2307–2327. [CrossRef]

24. Peng, M.; Xiang, L.; Wang, S. Semiparametric regression analysis of clustered survival data with semi-competing risks. *Comput. Stat. Data Anal.* **2018**, *124*, 53–70. [CrossRef]

25. Huang, X.-W.; Wang, W.; Emura, T. A copula-based Markov chain model for serially dependent event times with a dependent terminal event. *Jpn. J. Stat. Data Sci.* **2020**, *4*, 917–951. [CrossRef]

26. Shih, J.H. Common Mean. Copula: Bivariate Common Mean Vector under Copula Models; CRAN. 2022. Available online: https://CRAN.R-project.org/package=CommonMean.Copula (accessed on 8 January 2022).

27. Berkey, C.S.; Hoaglin, D.C.; Antczak-bouckoms, A.; Mosteller, F.; Colditz, G.A. Meta-analysis of multiple outcomes by regression with random effect. *Stat. Med.* **1998**, *17*, 2537–2550. [CrossRef]

28. Shinozaki, N. A note on estimating the common mean of k normal distributions and the stein problem. *Commun. Stat.-Theory Methods* **1978**, *7*, 1421–1432. [CrossRef]

29. Malekzadeh, A.; Kharrati-Kopaei, M. Inferences on the common mean of several normal populations under hetero-scedasticity. *Comput. Stat.* **2018**, *33*, 1367–1384. [CrossRef]

30. Gasparrini, A. Mvmeta: Multivariate and Univariate Meta-Analysis and Meta-Regression; CRAN. 2019. Available online: https://CRAN.R-project.org/package=mvmeta (accessed on 8 January 2022).

31. Nelsen, R. An Introduction to Copulas. *Technometrics. Lett.* **2000**, *42*, 317. [CrossRef]

32. Durante, F.; Sempi, C. *Principles of Copula Theory*; CRC/Chapman & Hall: Boca Raton, FL, USA, 2016.

33. Morgenstern, D. Einfache Beispiele zweidimensionaler Verteilungen. *Mitt. Für Math. Stat.* **1956**, *8*, 34–235.

34. Bairamov, I.G.; Kotz, S.; Bekci, M. New generalized Farlie-Gumbel-Morgenstern distributions and concomitants of order statistics. *J. Appl. Stat.* **2001**, *28*, 521–536. [CrossRef]

35. Bairamov, I.; Kotz, S. Dependence structure and symmetry of Hunag-Kotz FGM distributions and their extensions. *Metrika* **2002**, *56*, 55–72. [CrossRef]

36. Amini, M.; Jabbari, H.; Borzadaran, G.R.M. Aspects of Dependence in Generalized Farlie-Gumbel-Morgenstern Distributions. *Commun. Stat.-Simul. Comput.* **2011**, *40*, 1192–1205. [CrossRef]

37. Domma, F.; Giordano, S. A copula-based approach to account for dependence in stress-strength models. *Stat. Pap.* **2012**, *54*, 807–826. [CrossRef]

38. Chesneau, C. A new two-dimensional relation copula inspiring generalized version of the Farlie-Gumbel-Morgenstern copula. *Res. Commun. Math. Math. Sci.* **2021**, *13*, 99–128.

39. Clayton, D.G. A model for association in bivariate life tables and its application in epidemiological studies of familial tendency in chronic disease incidence. *Biometrika* **1978**, *65*, 141–151. [CrossRef]
40. Duchateau, L.; Janssen, P. *The Frailty Model*; Springer: New York, NY, USA, 2007.
41. Gumbel, E.J. Distributions de valeurs extrêmes en plusieurs dimensions. *Publ. Inst. Statist. Univ. Paris* **1960**, *9*, 171–173.
42. Frank, M.J. On the simultaneous associativity of F(x, y) and x+y−F(x, y). *Aequ. Math.* **1979**, *19*, 194–226. [CrossRef]
43. Sklar, A. Fonctions de répartition à n dimensions et leurs marges. *Publ. Inst. Statist. Univ. Paris* **1959**, *8*, 229–231.
44. Schucany, W.R.; Parr, W.C.; Boyer, J.E. Correlation structure in Farlie-Gumbel-Morgenstern distributions. *Biometrika* **1978**, *65*, 650–653. [CrossRef]
45. Nelsen, R.B. Dependence and Order in Families of Archimedean Copulas. *J. Multivar. Anal.* **1997**, *60*, 111–122. [CrossRef]
46. Bradley, R.A.; Gart, J.J. The asymptotic properties of ML estimators when sampling from associated populations. *Biometrika* **1962**, *49*, 205–214. [CrossRef]
47. Emura, T.; Hu, Y.H.; Konno, Y. Asymptotic inference for maximum likelihood estimators under the special exponential family with double-truncation. *Stat. Pap.* **2017**, *58*, 877–909. [CrossRef]
48. Shih, J.H. Copula-Based Statistical Inferences for a Common Mean Vector and Correlation Ratios Using Bivariate Data. Ph.D. Thesis, National Central University Library, Taoyuan, Taiwan, 2020. Available online: https://etd.lib.nctu.edu.tw/cgi-bin/gs32/ncugsweb.cgi/ccd=GLZeNP/record?r1=2&h1=2 (accessed on 3 October 2020).
49. Shao, J. *Mathematical Statistics*; Springer: New York, NY, USA, 2003.
50. Van der Vaart, A.W. *Asymptotic Statistics*; Cambridge University Press: Cambridge, UK, 1998.
51. Lehmann, E.L.; Casella, G. *Theory of Point Estimation*, 2nd ed.; Springer: New York, NY, USA, 1998.
52. Kontopantelis, E.; Reeves, D. Performance of statistical methods for meta-analysis when true study effects are non-normally distributed: A simulation study. *Stat. Methods Med Res.* **2010**, *21*, 409–426. [CrossRef]
53. Jackson, D.; White, I.R.; Riley, R.D. A matrix-based method of moments for fitting the multivariate random effects model for meta-analysis and meta-regression. *Biom. J.* **2013**, *55*, 231–245. [CrossRef]
54. Schepsmeier, U.; Stöber, J. Derivatives and Fisher information of bivariate copulas. *Stat. Pap.* **2013**, *55*, 525–542. [CrossRef]
55. Oakes, D. A Model for Association in Bivariate Survival Data. *J. R. Stat. Soc. Ser. B Stat. Methodol.* **1982**, *44*, 414–422. [CrossRef]
56. Nakatochi, M.; Kanai, M.; Nakayama, A.; Hishida, A.; Kawamura, Y.; Ichihara, S.; Matsuo, H. Genome-wide me-ta-analysis identifies multiple novel loci associated with serum uric acid levels in Japanese individuals. *Commun. Biol.* **2019**, *2*, 1–10. [CrossRef]
57. Taketomi, N.; Konno, Y.; Chang, Y.-T.; Emura, T. A Meta-Analysis for Simultaneously Estimating Individual Means with Shrinkage, Isotonic Regression and Pretests. *Axioms* **2021**, *10*, 267. [CrossRef]
58. Taketomi, N.; Michimae, H.; Chang, Y.-T.; Emura, T. meta.shrinkage: An R package for meta-analyses for simultaneously estimating individual means. *Algorithms*, 2022; *accepted*.