ORİJİNAL ARAŞTIRMA ORIGINAL RESEARCH

A Survival Copula Based Goodness-of-fit Testing for the Right-censored Case with Hazard Scenario

Sağdan Sansürlü Vaka İçin Sağkalım Kopula Tabanlı Uyum İyiliği Testi ve Hazard Senaryosu

Ece GÖRCEĞİZ[a], Burcu HÜDAVERDİ[a]

[a]Dokuz Eylül University Faculty of Sciences, Department of Statistics, İzmir, TURKEY

ABSTRACT Objective: Modeling dependence structure of a bivariate survival data is one of the main issues in biomedical studies. Survival copula deals with such a lifetime data and is used for modeling and understanding the distributional structure. In this study, we consider modeling and analysing the bivariate survival data in the presence of right censoring using Archimedean copula functions. In addition, the Hazard Scenario is modeled to assess the maximum threat expected from data sets examined in a particular area. So, this study has two main objectives. First one is to estimate the distributional structure of bivariate uncensored or right-censored survival data using bivariate survival Archimedean copula. Second one is to evaluate the joint risk of this bivariate data using Hazard Scenario approach after this distributional assumption. **Material and Methods:** We use Emura et al.(2010) goodness-of -fit testing procedure for the model selection. First, we examine the heart transplant data and model the dependence structure between waiting time for transplant and post-transplant survival time to see the co-movements of these variables. Second, we examine the diabetic retinopathy data and model the dependence between the survival times of the two eyes of the same patient in case of laser photocoagulation treatment. Finally, we use the Survival Hazard Scenario approach to evaluate the probability of exceeding some critical layers for the data. Some graphs based on λ-functions and Kendall functions are also presented for visual comparison. **Results:** Based on the goodness-of-fit procedure applied to both of the datasets for the model selection, some parametric models are determined. These models help to understand the joint behaviour of the data. For both of the data sets, Frank copula which has symmetric dependence structure is selected. The data sets have symmetric co-movements at the tails and around the center. For the heart transplant data, it can be concluded if the patient is strong enough to wait for the heart transplant process, the survival time after transplantation is also longer. For the retinopathy data, the survival times of two eyes move together after the laser photocoagulation treatment. After the determination of the bivariate Archimedean copula of the data sets, a Survival Hazard Scenario which is a probabilistic consistent framework is used to obtain some joint risk layers. These layers help us to determine the large values of the variables which are associated with high risk conditions. **Conclusion:** The symmetric dependence structure is observed for both of the data set. Also, the data pairs which exceed the critical layers are also obtained using survival Kendall Hazard Scenario which provides valuable resulting for assessing the probability of threatening events.

Keywords: Survival Hazard Scenario; right censoring;
bivariate survival data; Archimedean copula;
survival copula

ÖZET Amaç: İki değişkenli sağkalım verilerinin bağımlılık yapısının modellenmesi, biyomedikal çalışmaların ana konularından biridir. Böyle bir yaşam zamanı verisiyle ilgilenen sağkalım kopula modelleri bu verilerin dağılım yapısını modellemek ve anlamak için kullanılır. Bu çalışmada, Archimedean kopulalar kullanılarak sağdan sansürlü veri yapısı modellenmiş ve analiz edilmiştir. Ayrıca, belirli bir alanda incelenen veri setlerinden beklenen maksimum tehdidi değerlendirmek amacıyla Hazard Senaryosu modellenmiştir. Dolayısıyla, bu çalışmanın iki temel amacı vardır. Birincisi, iki değişkenli sağkalım Archimedean kopula kullanılarak sansürsüz ve sağdan sansürlü sağkalım verilerin dağılım yapısını tahmin etmektir. İkincisi, bu dağılımsal varsayımdan sonra Hazard Senaryosu yaklaşımını kullanarak bu iki değişkenli verilerin ortak riskini değerlendirmektir. **Gereç ve Yöntemler:** Model seçimi için Emura vd. (2010) uyum iyiliği prosedürü kullanılmıştır. ilk olarak, kalp nakli verileri incelenmiş ve bu değişkenlerin ortak hareketlerini görmek için nakil için bekleme süresi ile nakil sonrası hayatta kalma süresi arasındaki bağımlılık yapısını modellenmiştir. İkinci olarak, diyabetik retinopati verilerini incelenmiş ve lazer fotokoagülasyon tedavisi durumunda aynı hastanın iki gözünün yaşam süreleri arasındaki bağımlılığı modellenmiştir. Son olarak, veriler için bazı kritik seviyeleri aşma olasılığı değerlendirmek üzere sağkalım Hazard Senaryosu yaklaşımı kullanılmıştır. λ-fonksiyonlarına ve Kendall fonksiyonlarına dayanan bazı grafikler de görsel karşılaştırma için sunulmuştur. **Bulgular:** Model seçimi için her iki veri kümesine uygulanan uyum iyiliği sürecine dayanarak, bazı parametrik modeller belirlenmiştir. Bu modeller verilerin ortak davranışını anlamaya yardımcı olur. Her iki veri seti için de, simetrik bağımlılık yapısına sahip olan Frank kopula seçilmiştir.. Veri yapıları merkez etrafında ve kuyrukta birlikte hareket etmektedir. Kalp nakli verisi için hastanın kalp nakli sürecini bekleyecek kadar güçlü olması durumunda, nakil sonrası hayatta kalma süresi de daha uzundur yorumu yapılabilir. Retinopati verileri için lazer fotokoagülasyon tedavisinden sonra iki gözün sağkalım süresi de birlikte hareket etmektedir. İki değişkenli Archimedean kopula yapısının belirlenmesinin ardından, bazı ortak risk katmanları elde etmek için olasılıksal bir yapı olan iki değişkenli sağkalım Hazard Senaryosu kullanılmıştır. Bu risk seviyeleri, değişkenlerin büyük değerlerinin riskli koşullarla ilişkili olduğu değerleri belirlememize yardımcı olmaktadır. **Sonuç:** Simetrik bağımlılık yapısı her iki veri setinde de gözlenmektedir. Ayrıca, kritik seviyeleri aşan veri çiftleri, riskli olayların olasılığını değerlendirmek için önemli sonuçlar sağlayan Kendall sağkalım Hazard Senaryosu kullanılarak da elde edilmiştir.

Anahtar kelimeler: Sağkalım Hazard Senaryosu; sağdan sansür;
iki değişkenli sağkalım verisi;
Archimedean kopula; sağkalım kopulası

Modeling bivariate and multivariate survival data recently has a growing interest. In the biomedical area, the dependence structure of the bivariate survival data has been studied by many researchers. Copulas are key tools to analyse the dependence structure. A bivariate survival function can be formed by the marginal survival functions and their bivariate copula. Since a copula is a great deal of flexibility in modeling bivariate survival data, it provides an effective approach for understanding and modeling the dependent random variables and so the dependence structure.

Copula models examine the joint distribution in two ways: the dependence between the variables and the marginal distributions of the individual variables. Hence the dependence structure may also be obtained separately from the marginal distributions and the copulas. One-parameter family of copulas is preferred for modeling dependence. The Archimedean family of copulas is an important family of copulas. They can be reduced to a single univariate distribution function called as Kendall distribution function, $K(.)$. In this study, we examine the goodness-of-fit (GoF) testing procedure of Archimedean class for bivariate survival data under uncensored and censored cases given by Emura et al.[1] A model selection procedure which is based on GoF techniques for censored data in bivariate survival models was first suggested by Wang and Wells.[2]

Let (X,Y) be a random pair of bivariate survival time with bivariate survival function $S(x,y)$ so the copula can be expressed as

$$S(x,y) = C(S(x),S(y))$$

where $S(x)$ and $S(y)$ are marginal survival functions and $C(u,v):[0,1]^2 \rightarrow [0,1]$ is the copula function. For Archimedean copulas, a generator function φ is continuous and strictly decreasing $\varphi_\alpha(.):[0,1] \rightarrow [0,\infty]$ with a dependence parameter α. An Archimedean copula has a form in terms of generator function φ such as

$$C(u,v) = \varphi^{[-1]}(\varphi_\alpha(u) + \varphi_\alpha(v)).$$

See, Nelsen[3] for more details.

In this study, we deal with three members which are Gumbel, Frank and Clayton of the Archimedean copula family. They have different dependence characteristics. Clayton copula has strong left tail dependence and relatively weak right tail dependence. When α approaches to zero, Clayton copula becomes independence copula, and when α approaches to infinity, Clayton copula reaches the Frèchet upper bound, but it reaches the Frèchet lower bound for no value.

Frank copula allows negative dependence between the marginals. It interpolates between perfectly negative dependence and perfectly positive dependence. Frank copula has symmetric dependence in both tails. If α approaches to zero, Frank copula approaches to the independence copula; if α approaches to infinity, Frank copula becomes the Frèchet upper bound; similarly, if α approaches to negative infinity, Frank copula becomes the Frèchet lower bound, which equal to the perfectly negative dependence.

Gumbel copula allows perfectly positive dependence, but does not allow negative dependence. The Gumbel copula shows strong right tail dependence and relatively weak left tail dependence. It is an appropriate choice when outcomes are known to be strongly correlated at high values but less correlated at low values. If α is equal to 1, Gumbel copula is an independent copula. If α approaches to infinity, the limit of Gumbel copula reaches the Frèchet upper bound.

The local odds ratio function is proposed by Oakes[4] given as,

$$\theta^*(x,y) = \frac{\partial^2 P(X>x,Y>y)/\partial x \partial y P(X>x,Y>Y)}{\partial P(X>x,Y>y)/\partial x P(X>x,Y>y)/\partial y}$$

$$= \frac{P(\Delta_{ij} = 1 \mid \tilde{X}_{ij} > x, \tilde{Y}_{ij} > y)}{P(\Delta_{ij} = 0 \mid \tilde{X}_{ij} > x, \tilde{Y}_{ij} > y)}$$

where $\tilde{X}_{ij} = X_i \wedge X_j$ and $\tilde{Y}_{ij} = Y_i \wedge Y_j$, and $a \wedge b = \min(a,b)$. The procedures work well with survival censored data. Emura et al.[5] indicated that a cross-ratio function can characterize the results of inferential procedures for censored and uncensored data. $\theta^*(x, y)$ is a bivariate function which measures the local dependence when the pair (X, Y) is independent at (x, y). Also, $\theta^*(x, y) = \theta\{F(x, y)\}$ for the Archimedean copula class. $\theta\{F(x, y)\}$ is obtained as

$$\theta_\alpha(v) = -v \left\{ \frac{\left[\phi_\alpha''(v)\right]}{\left[\phi_\alpha'(v)\right]} \right\},$$

Genest and Rivest[6] indicated that $K(v)$ is related to $\phi_\alpha(v)$ through the differential equation

$$\lambda(v) = v - K(v) = \frac{\phi_\alpha(v)}{\phi_\alpha'(v)}$$

where $\phi_\alpha'(v) = \partial\phi_\alpha(v)/\partial v$ and $v = S(X, Y)$. See also, Wang and Wells[2]. The function $\phi_\alpha(v)$ can be obtained by $K(v)$ where $K(v) \equiv P\{F(X, Y) \le v\}$. $\lambda(v)$ is related with the generator function $\phi_\alpha(v)$ and determines the dependence structure of the Archimedean copula class. Table 1 gives the generator and the lambda functions of some Archimedean copulas.

TABLE 1: Some Archimedean Copulas.

	ϕ	λ	τ
Clayton	$\left(S\left(\tilde{X}_{ij}, \tilde{Y}_{ij}\right)^{-\alpha} - 1\right)/\alpha$	$-S\left(\tilde{X}_{ij}, \tilde{Y}_{ij}\right)\left(1 - S\left(\tilde{X}_{ij}, \tilde{Y}_{ij}\right)^\alpha\right)/\alpha$	$\alpha/(\alpha+2)$
Frank	$\log\left(\dfrac{1 - \exp(-\alpha)}{1 - \exp\left(-\alpha S\left(\tilde{X}_{ij}, \tilde{Y}_{ij}\right)\right)}\right)$	$-\dfrac{1 - \exp\left(-\alpha S\left(\tilde{X}_{ij}, \tilde{Y}_{ij}\right)\right)}{\alpha \exp\left(-\alpha S\left(\tilde{X}_{ij}, \tilde{Y}_{ij}\right)\right)} \log\left(\dfrac{1 - \exp(-\alpha)}{1 - \exp\left(-\alpha S\left(\tilde{X}_{ij}, \tilde{Y}_{ij}\right)\right)}\right)$	$1 + 4\{D_1(\alpha) - 1\}/\alpha$
Gumbel	$\left\{-\log\left(S(\tilde{X}_{ij}, \tilde{Y}_{ij})\right)\right\}^{\alpha+1}$	$S\left(\tilde{X}_{ij}, \tilde{Y}_{ij}\right)\log\left(S\left(\tilde{X}_{ij}, \tilde{Y}_{ij}\right)\right)/(\alpha+1)$	$\alpha/(\alpha+1)$

NOTE: α is the dependence parameter.

* D_1 is the Debye function of order 1, $D_1(\alpha) = \int_0^\alpha \left\{t/\alpha\left(e^t - 1\right)\right\}dt$. [6]

There are two main contributions of this study to the existing literature. First, we use goodness-of-fit procedure in the case of uncensored and right-censored data for some Archimedean copula models and obtain their test statistics. These Archimedean copula models are Frank, Gumbel and Clayton which have different types of dependence characteristics such as, symmetric dependence, right-tail dependence and left tail dependence, respectively. The goodness-of-fit test procedure is applied to check which copula fits best to the given data. Some graphical tools and numerical techniques are used in this process to select a suitable and adequate model. The goodness-of-fit tests constructed as $H_0 : C \in C_\theta$ versus $H_1 : C \notin C_\theta$ where C is an unknown copula and C_0 is the parametric copula. Graphical methods, error statistics, and goodness-of-fit statistics are used to measure the adequacy of the hypotheses. Second, we employ Survival Hazard scenario

approach based on copula models to determine the critical layers for the bivariate survival data. The critical layers show the high risk thresholds for the pair of the data. In Materials and Methods section, the testing procedure is given for both the censored and the uncensored data. The survival hazard scenario in terms of survival Kendall distribution is also given in this section. In Results section, all the given procedures are applied to the heart transplant data and diabetic retinopathy data sets. Conclusion section is also given.

MATERIAL AND METHODS

In this section, we deal with the uncensored data firstly, then we modify the model for the right censored data. The main hypothesis is given by

$$H_0 : C(u,v) = \phi_\alpha^{-1}\left[\phi_\alpha(u) + \phi_\alpha(v)\right] \text{ for some } \alpha \in \Re,$$

where the alternative hypothesis is any other copula.

The GOF procedure for the uncensored data

Let $(X_i, Y_i), i = 1, \dots, n$ be the uncensored data, and Δ_{ij} be the concordance indicator, then

$$U_k(\alpha) = \sum_{i<j} W_k(\tilde{X}_{ij}, \tilde{Y}_{ij})\left[\Delta_{ij} - E\left[\Delta_{ij} \mid \tilde{X}_{ij} = x, \tilde{Y}_{ij} = y\right]\right]$$

$$U_k(\alpha) = \sum_{i<j} W_k(\tilde{X}_{ij}, \tilde{Y}_{ij})\left[\Delta_{ij} - \frac{\theta_\alpha\{\hat{S}(\tilde{X}_{ij}, \tilde{Y}_{ij})\}}{\theta_\alpha\{\hat{S}(\tilde{X}_{ij}, \tilde{Y}_{ij})\}+1}\right]$$

where $W_k(\tilde{X}_{ij}, \tilde{Y}_{ij})$ is weight function and $\hat{S}(x,y)$ is an estimator of $S(x,y)$. $\hat{S}(x,y)$ is called 'empirical survival function' obtained as $\hat{S}(x,y) = \frac{1}{n}\sum_i I(X_i \geq x, Y_i \geq y)$. In goodness-of-fit procedure, two weight functions are used to obtain $\hat{\alpha}_k$ by solving $U_k(\alpha) = 0$ for $k = 1, 2$. If H_0 is true, then the statistic $n^{1/2}\{\hat{\alpha}_1 - \hat{\alpha}_2\}$ converges to a normal distribution with zero mean. As Emura et al.[1] indicated, the power of the test depends on the weight functions.

THE WEIGHT FUNCTIONS

The unweighted and weighted estimators of the association parameter were compared by Shih[7] and the conclusion was that if the proposed model has a good fit, the difference of these two estimates converges to zero. Using the idea of the likelihood approach of Clayton[8], the estimating function can be written in terms of $U_k(\alpha)$. Let Ψ be the set of grid points, then

$$\psi = \left\{(x,y)\left|\sum_{i=1}^n I(X_i = x, Y_i \geq y) = 1, \sum_{i=1}^n I(X_i \geq x, Yi = y) = 1\right.\right\}$$

Also, let $D(x,y)$ be the number of observed failures defined as

$$D(x,y) = \sum_i I(X_i = x, Y_i = y)$$

and let $R(x,y)$ measure the number of risk defined as

$$R(x,y) = r = \sum_{i=1} I(X_i \geq x, Y_i \geq y),$$

then, $D(x, y)$ is distributed Bernoulli with the success probability such as

$$P\{D(x, y) = 1 | R(x, y) = r, (x, y) \in \psi\} = \frac{\theta_\alpha \{S(x, y)\}}{r - 1 + \theta_\alpha \{S(x, y)\}}$$

By using conditional probability, the likelihood function is

$$L(\alpha) = \prod_{(x,y) \in \psi} \left[\frac{\theta_\alpha \{S(x, y)\}}{R(x, y) - 1 + \theta_\alpha \{S(x, y)\}} \right]^{D(x,y)} \times \left[\frac{R(x, y) - 1}{R(x, y) - 1 + \theta \{S(x, y)\}} \right]^{1 - D(x,y)}$$

with all points in ψ under the independence assumption among the grids.

Then, the estimating function is

$$U_1(\alpha) = \sum_{i<j} \frac{\dot\theta_\alpha \{\hat{S}(X_{ij}^0, Y_{ij}^0)\} \left[\theta_\alpha \{\hat{S}(X_{ij}^0, Y_{ij}^0)\} + 1 \right]}{\theta_\alpha \{\hat{S}(X_{ij}^0, Y_{ij}^0)\} \left[R_{ij} - 1 + \theta_\alpha \{\hat{S}(X_{ij}^0, Y_{ij}^0)\} \right]} \left[\Delta_{ij} - \frac{\theta_\alpha \{\hat{S}(X_{ij}^0, Y_{ij}^0)\}}{\theta_\alpha \{\hat{S}(X_{ij}^0, Y_{ij}^0)\} + 1} \right]$$

$$= \sum_{i<j} W_1(X_{ij}^0, Y_{ij}^0) \left[\Delta_{ij} - \frac{\theta_\alpha \{\hat{S}(X_{ij}^0, Y_{ij}^0)\}}{\theta_\alpha \{\hat{S}(X_{ij}^0, Y_{ij}^0)\} + 1} \right],$$

where $R_{ij} = R(X_{ij}^0, Y_{ij}^0) = n\hat{S}(X_{ij}^0, Y_{ij}^0)$ and $\dot\theta_\alpha \{\hat{S}(X_{ij}^0, Y_{ij}^0)\} = \partial\theta_\alpha \{\hat{S}(X_{ij}^0, Y_{ij}^0)\} / \partial\alpha$. The second estimating function is

$$U_2(\alpha) = \sum_{i<j} \left[\Delta_{ij} - \frac{\theta_\alpha \{\hat{S}(X_{ij}^0, Y_{ij}^0)\}}{\theta_\alpha \{\hat{S}(X_{ij}^0, Y_{ij}^0)\} + 1} \right].$$

By solving $U_1(\alpha) = 0$ and $U_2(\alpha) = 0$, we find $\hat\alpha_k$ for $k = 1, 2$.

We obtain the U statistics for the given Archimedean copula functions.

For Frank copula,

$$U_1(\alpha) = \sum_{i<j} \frac{\hat{S}(X_{ij}^0, Y_{ij}^0) / (1 - \exp(-\hat{S}(X_{ij}^0, Y_{ij}^0)\alpha)) - \hat{S}(X_{ij}^0, Y_{ij}^0)^2 \alpha / (1 - \exp(-\hat{S}(X_{ij}^0, Y_{ij}^0)\alpha))^2 \exp(-\hat{S}(X_{ij}^0, Y_{ij}^0)a) \left[(\hat{S}(X_{ij}^0, Y_{ij}^0)\alpha) / (1 - \exp(-\alpha\hat{S}(X_{ij}^0, Y_{ij}^0))) + 1 \right]}{\left[(\hat{S}(X_{ij}^0, Y_{ij}^0)\alpha) / (1 - \exp(-\alpha\hat{S}(X_{ij}^0, Y_{ij}^0))) \right] \left[R_{ij} - 1 + \left[(\hat{S}(X_{ij}^0, Y_{ij}^0)\alpha) / (1 - \exp(-\alpha\hat{S}(X_{ij}^0, Y_{ij}^0))) \right] \right]}$$

$$\times \left[\Delta_{ij} - \frac{\left[(\hat{S}(X_{ij}^0, Y_{ij}^0)\alpha) / (1 - \exp(-\alpha\hat{S}(X_{ij}^0, Y_{ij}^0))) \right]}{\left[(\hat{S}(X_{ij}^0, Y_{ij}^0)\alpha) / (1 - \exp(-\alpha\hat{S}(X_{ij}^0, Y_{ij}^0))) + 1 \right]} \right]$$

$$U_2(\alpha) = \sum_{i<j} \left[\Delta_{ij} - \frac{\left[(\hat{S}(X_{ij}^0, Y_{ij}^0)\alpha) / (1 - \exp(-\alpha\hat{S}(X_{ij}^0, Y_{ij}^0))) \right]}{\left[(\hat{S}(X_{ij}^0, Y_{ij}^0)\alpha) / (1 - \exp(-\alpha\hat{S}(X_{ij}^0, Y_{ij}^0))) + 1 \right]} \right]$$

For Gumbel copula,

$$U_1(\alpha) = \sum_{i<j} \frac{2\log \hat{S}(X_{ij}^0, Y_{ij}^0) - \alpha}{\{\log \hat{S}(X_{ij}^0, Y_{ij}^0) - \alpha\}\{\alpha - R_{ij} \log \hat{S}(X_{ij}^0, Y_{ij}^0)\}} \left[\Delta_{ij} - \frac{\log \hat{S}(X_{ij}^0, Y_{ij}^0) - \alpha}{2\log \hat{S}(X_{ij}^0, Y_{ij}^0) - \alpha} \right]$$

$$U_2(\alpha) = \sum_{i<j} \left[\Delta_{ij} - \frac{\log \hat{S}(X_{ij}^0, Y_{ij}^0) - \alpha}{2\log \hat{S}(X_{ij}^0, Y_{ij}^0) - \alpha} \right]$$

For Clayton copula,

$$U_1(\alpha) = \sum_{i<j} \frac{[(\alpha+1)+1]}{(\alpha+1)\hat{S}(\tilde{X}_{ij},\tilde{Y}_{ij})}\left[\Delta_{ij} - \frac{(\alpha+1)}{[(\alpha+1)+1]}\right]$$

$$U_2(\alpha) = \sum_{i<j}\left[\Delta_{ij} - \frac{(\alpha+1)}{[(\alpha+1)+1]}\right]$$

For, the asymptotic distribution of $n^{1/2}(\hat{\alpha}_1 - \hat{\alpha}_2)$ and $n^{1/2}(\log\hat{\alpha}_1 - \log\hat{\alpha}_2)$ approximates to $N(0,\sigma^2)$, where $\sigma^2 = 4E\left[h\{(X_1,Y_1),(X_2,Y_2)\}h\{(X_1,Y_1),(X_3,Y_3)\}\right]$. And also $(\log\hat{\alpha}_1 - \log\hat{\alpha}_2)$ converges zero in probability, so that $U_1(\alpha)$ and $U_2(\alpha)$ involve the estimator.

Lemma: Under the correct model and regularity conditions

$$n\binom{n}{2}^{-1}U_1(\alpha) = \binom{n}{2}^{-1}\tilde{U}_1(\alpha) + O_p(1), \qquad \binom{n}{2}^{-1}U_2(\alpha) = \binom{n}{2}^{-1}\tilde{U}_2(\alpha) + O_p(1),$$

where $O_p(1)$ is uniform in α.

Lemma: When $\gamma = \log\alpha$ and $\hat{\gamma}_k = \log\hat{\alpha}_k$

$$n^{1/2}(\hat{\gamma}_1 - \hat{\gamma}_2) = n^{1/2}\binom{n}{2}^{-1}\sum_{i<j}h\{(X_i,Y_i),(X_j,Y_j)\} + O_p(1),$$

where h is a symmetric function and

$$h\{(X_i,Y_i),(X_j,Y_j)\} \equiv \frac{1}{\alpha}\left(\frac{\dot{\theta}_\alpha\{S(\tilde{X}_{ij},\tilde{Y}_{ij})\}[\theta_\alpha\{S(\tilde{X}_{ij},\tilde{Y}_{ij})\}+1]}{A_L\theta_\alpha\{S(\tilde{X}_{ij},\tilde{Y}_{ij})\}S(\tilde{X}_{ij},\tilde{Y}_{ij})} - \frac{1}{A}\right)\left[\Delta_{ij} - \frac{\theta_\alpha\{S(\tilde{X}_{ij},\tilde{Y}_{ij})\}}{\theta_\alpha\{S(\tilde{X}_{ij},\tilde{Y}_{ij})\}+1}\right],$$

where $A \equiv E\left(\frac{\dot{\theta}_\alpha\{S(\tilde{X}_{12},\tilde{Y}_{12})\}}{[\theta_\alpha\{S(\tilde{X}_{12},\tilde{Y}_{12})\}+1]^2}\right)$ and $A_L \equiv E\left(\frac{\dot{\theta}_\alpha[\{S(\tilde{X}_{12},\tilde{Y}_{12})\}]^2}{\theta_\alpha\{S(\tilde{X}_{12},\tilde{Y}_{12})\}[\theta_\alpha\{S(\tilde{X}_{12},\tilde{Y}_{12})\}+1]}\right)$.

THE TESTING PROCEDURE

When $Z_0 = |\hat{\gamma}_1 - \hat{\gamma}_2|/\hat{\sigma} > Z_{1-\alpha/2}$ under α-significance level where $Z_{1-\alpha/2}$ is the *p*th percentile of the standard normal distribution, we reject H_0. In Archimedean copula models, the derivation of the variance estimator is difficult and also it becomes more complex for right-censored data. Therefore, for the variance estimation, Emura et al.[1] proposed the jackknife method. The jackknife is a method which is one of the earliest resampling methods used to estimate the variance and bias of a large population. The method is defined as

$$\sigma_{jack}^2 = \frac{n-1}{n}\sum_{i=1}^{n}\left\{\hat{\gamma}_1^{(-i)} - \hat{\gamma}_2^{(-i)} - (\hat{\gamma}_1^{(\cdot)} - \hat{\gamma}_2^{(\cdot)})\right\}^2,$$

where $\hat{\gamma}^{(-i)}$ is the estimator after deleting the ith observation and $\hat{\gamma}_1^{(\cdot)} - \hat{\gamma}_2^{(\cdot)} = \frac{1}{n}\sum_{i=1}^{n}\hat{\gamma}_1^{(-i)} - \hat{\gamma}_2^{(-i)}$

RIGHT-CENSORED CASE

Let (A_i, B_i) be the independently identically distributed bivariate censoring variables. (A_i, B_i) and (X_i, Y_i) are independent of each other. Through the right censoring process, it is applied that, $\tilde{Y}_i = \min(Y_i, B_i)$, $\delta_i^x = I(X_i \le A_i)$ and $\delta_i^y = I(Y_i \le B_i)$. The order of X_i and X_j is known if and only if $\tilde{X}_{ij} \le \tilde{A}_{ij}$, where $\tilde{X}_{ij} = X_i \wedge X_j$ and $\tilde{A}_{ij} = A_i \wedge A_j$. Similarly, the order of Y_i and Y_j is known if and only if $\tilde{Y}_{ij} \le \tilde{B}_{ij}$. Let Z_{ij} indicate whether the ordering relationship is certain or not. Then, then U- statistic is

$$U_k(\alpha) = \sum_{i<j} Z_{ij} W_k(\tilde{X}_{ij}, \tilde{Y}_{ij}, \alpha, \hat{S}) \left[\Delta_{ij} - \frac{\theta_\alpha \left\{ \hat{S}(\tilde{X}_{ij}, \tilde{Y}_{ij}) \right\}}{\theta_\alpha \left\{ \hat{S}(\tilde{X}_{ij}, \tilde{Y}_{ij}) \right\} + 1} \right] \quad k = (1,2) \; ,$$

where $\hat{S}(x,y)$ is an estimator of $S(x,y)$ and $\Delta_{ij} = I\left\{ (\hat{X}_i - \hat{X}_j)(\hat{Y}_i - \hat{Y}_j) > 0 \right\}$. The weight function is

$$W_1(\tilde{X}_{ij}, \tilde{Y}_{ij}, \alpha, S) = \frac{\dot{\theta}_\alpha \left\{ S(\tilde{X}_{ij}, \tilde{Y}_{ij}) \right\} \left[\theta_\alpha \left\{ S(\tilde{X}_{ij}, \tilde{Y}_{ij}) \right\} + 1 \right]}{\theta_\alpha \left\{ S(\tilde{X}_{ij}, \tilde{Y}_{ij}) \right\} \left[R_{ij} - 1 + \theta_\alpha \left\{ S(\tilde{X}_{ij}, \tilde{Y}_{ij}) \right\} \right]}$$

where $W_2(\tilde{X}_{ij}, \tilde{Y}_{ij}, \alpha, S) = 1$ and $R_{ij} = \sum_{i=1}^{n} I(\hat{X}_I \ge \tilde{X}_{ij}, \hat{Y}_I \ge \tilde{Y}_{ij})$.

To solve $\hat{\alpha}_k$, the following equation is used;

$$\sum_{i<j} Z_{ij} W_k(\tilde{X}_{ij}, \tilde{Y}_{ij}, \alpha, \hat{S}^{(l-1)}) \left[\Delta_{ij} - \frac{\theta_\alpha \left\{ \hat{S}^{(l-1)}(\tilde{X}_{ij}, \tilde{Y}_{ij}) \right\}}{\theta_\alpha \left\{ \hat{S}^{(l-1)}(\tilde{X}_{ij}, \tilde{Y}_{ij}) \right\} + 1} \right] = 0 \qquad l = 1, 2, \dots$$

where $\hat{S}^{(l)}$ shows the estimated value in the l-step and also defined as $\hat{S}_\alpha^0(x,y)$.

As in the uncensored case, the results of asymptotic normality are valid and

$$\tilde{U}_2(\alpha) = \sum_{i<j} Z_{ij} \left[\Delta_{ij} - \frac{\theta_\alpha \left\{ S(\tilde{X}_{ij}, \tilde{Y}_{ij}) \right\}}{\theta_\alpha \left\{ S(\tilde{X}_{ij}, \tilde{Y}_{ij}) \right\} + 1} \right] ,$$

U -statistic approximate to $U_k(\alpha)$ as in the above $U_2(\alpha)$ function:

$$\tilde{U}_2(\alpha) = \sum_{i<j} Z_{ij} \left[\Delta_{ij} - \frac{\theta_\alpha \left\{ S(\tilde{X}_{ij}, \tilde{Y}_{ij}) \right\}}{\theta_\alpha \left\{ S(\tilde{X}_{ij}, \tilde{Y}_{ij}) \right\} + 1} \right] .$$

HAZARD SCENARIO

"Hazard" is a situation or property with the potential to construct harm and also is known as the potential for an accident with undesired results in process safety. Modeling the hazard scenario aims to evaluate the maximum threat expected from a studied event in a certain area. Hazard Scenarios should be multivariate and dependent. In Hazard Scenarios, several scenarios are tendered and comparisons are applied. Firstly we deal with some notation in order to explain the Hazard Scenario. Salvadori et al.[9] and Salvadori et al.[10] gave and discussed the general concept of the Hazard Scenario. Susam and Ucer[11] applied the Kendall scenario to the energy consumption data.

$S \subseteq R^d$ is defined as an upper set if, and only if, $x \in S$ and $y \geq x$, $y \in S$. The upper set handles the a practical reason of an event which can be classified as risky with "larger" realizations.

In this approach, X is a random vector defining the event of interest. For a level $\alpha \in (0,1)$, a Hazard Scenario is constructed on any Upper Set $S = S_\alpha \subseteq R^d$ relation

$$P(X \in S) = \alpha$$

holds. The dependence upon α will be eliminated when no complexity may occur to hold formula simple. Under an appropriate criterion, a Hazard Scenario can be constructed as a set which contains all the events x which are classified as "dangerous" events. When $x \in S$ and $y \geq x$, y could be classified as dangerous.

In a Hazard Scenario which can be defined as an Upper Set, generally the cases where the large values of the variables of interest are related to the 'dangerous' cases. Sometimes, small values of the variables may be dangerous also. In this case, the sign of the variables of interest is changed.

We use Kendall and Survival Kendall scenario to define the critical layers.

1. "Kendall" scenario S^K: Let $x \in R^d$ be a given event where $t = F(x)$ and let L_t be the level set crossing x. A bivariate Kendall Hazard Scenario is considered as the pair $(U, V) \in I^2$. Also, S_t^K is defined as the region "exceeding" the critical layer, and L_t and L_t represent a critical multivariate threshold.

2. "Survival Kendall" scenario $S^{\check{k}}$: Let $x \in R^d$ be a given event where $t = \overline{F}(x)$ and let $\overline{L}_t$ be the survival level set crossing x. A bivariate Survival Kendall Hazard Scenario is considered as the pair $(U, V) \in I^2$. Also, $S_t^{\check{k}}$ is defined as the region "exceeding" the survival critical layer, and $\overline{L}_t$ and $\overline{L}_t$ represent a critical multivariate threshold.

TABLE 2: Some Hazard Scenarios.

Scenarios	Shapes	Probabilities
Kendall	$S_t^K = \{y \in R^d : F(y) > t\} = \{y \in R^d : C(F_1(y_1),...,F_d(y_d)) > t\}$	$\alpha_u^K = \alpha_x^K = \alpha_t^K = P(X \in S_t^K) = 1 - K(t)$
Survival Kendall	$S_t^{\check{k}} = \{y \in R^d : \overline{F}(y) < t\} = \{y \in R^d : \hat{C}(\overline{F}_1(y_1),...,\overline{F}_d(y_d)) < t\}$	$\alpha_u^{\check{k}} = \alpha_x^{\check{k}} = \alpha_t^{\check{k}} = P(X \in S_t^{\check{k}}) = 1 - \check{K}(t) = \hat{K}(t)$

1. The Kendall case: Suppose that $u = F(x_1, x_2)$ is fixed and L_t be the critical layer crossing u which $t = C(u)$.

$$S_u^{\wedge} \subseteq S_t^K \subseteq S_u^v$$

and

$$\alpha_u^{\wedge} \leq \alpha_t^K \leq \alpha_u^v$$

α_u^v is defined as constant as u crosses over L_t, as well as α_t^K, but $\alpha_u^{\wedge}$ usually changes. We define S_t^K as unique for all events having L_t, the Hazard Scenario S_u^v's restate.

2. The Survival Kendall case: Suppose that $u = \bar{F}(x_1, x_2) \in I^d$ is fixed and let $\bar{L}_t$ be the survival critical layer crossing u which $t = \hat{C}(1-u)$.

$$S_u^{\wedge} \subseteq S_t^{\check{K}} \subseteq S_u^{v}$$

and

$$\alpha_u^{\wedge} \leq \alpha_t^{\check{K}} \leq \alpha_u^{v}$$

$\alpha_u^{\wedge}$ is defined as constant where u crosses over $\bar{L}_t$, as well as $\alpha_t^{\check{K}}$, but α_u^{v} usually changes. We define $S_t^{\check{K}}$ as unique for all events having $\bar{L}_t$, the Hazard Scenario $S_u^{\wedge}$'s restate.

The Survival Kendall case indicates the regions of bounded safe events, thus ensuring that non-hazardous events take bounded values of the variables.

For the Kendall Scenario probabilities

$$\begin{aligned}
\alpha_{u_1,u_2}^{K} &= 1 - K(t) \\
&= 1 - K\left(C(u_1,u_2)\right) \\
&= 1 - K\left(u_1 + u_2 - 1 + \alpha_{u_1,u_2}^{\wedge}\right)
\end{aligned}$$

For the Survival Kendall Scenario probabilities,

$$\begin{aligned}
\alpha_{u_1,u_2}^{\hat{K}} &= 1 - \check{K}(t) \\
&= 1 - \check{K}\left(\hat{C}(1-u_1, 1-u_2)\right) \\
&= 1 - \check{K}\left(\alpha_{u_1,u_2}^{\wedge}\right)
\end{aligned}$$

where $\check{K}(t)$ is the Survival Kendall distribution function and $\check{K}(t) = 1 - \hat{K}(t)$. For summary, Table 2 which presents the Hazard Scenarios and probabilities is given.

In the Failure Probability approach, let $T > 0$ be an arbitrary design lifetime and let $X_1, ..., X_T$ be the random vectors which defines the event in an investigation at times $1, ..., T$. Let $(S_1, ..., S_T)$ be Hazard Scenario vector and suppose that S_i^{c}'s are Hazard Scenario's complements namely their complements S_i^{c}'s could be labeled as "safe". For The Failure Probability Approach, it can be denoted as general formula;

$$\rho_T = 1 - P\left(X_1 \in S_1^{c}, ..., X_T \in S_T^{c}\right).$$

For the Hazard Scenario, the variables are iid, then

$$\rho_T = 1 - \prod_{j=1}^{T} P\left(X \in S^{c}\right) = 1 - (1-\alpha)^{T}.$$

Finally, for the Survival Kendall Scenario, the Failure Probability is

$$\rho_{\mathrm{T}}^{\acute{K}} = 1 - \grave{K}(t)^{\mathrm{T}}$$

for $t = \overline{F}(x_0)$.

RESULTS

Firstly, our data is obtained from heart transplant recipients from the Stanford heart transplant program.[12] We tackled 69 transplant data from 103 transplants and also our data is used for the uncensored case. Descriptive statistics of this data set is given in Table3.

TABLE 3: Descriptive Statistics of Heart Transplant Data.

Heart transplant	Mean	std	Var	Min	max	Median
Survival time	415.4	458.6698	210378	5	1799	207
Waiting time	38	50.10384	2510.394	1	310	26

We model the dependence structure between waiting time for transplant and post-transplant survival time to see the co-movements of these variables and how they affect each other.

We apply the goodness-of-fit procedure given in the previous section. Table 4 shows the goodness-of-fit results. The weighted and unweighted dependence parameters, test statistics and their p-values are shown in Table 4. For the Frank copula model, we obtain $\hat{\alpha}_1 = 2.1930472$ and $\hat{\alpha}_2 = 2.134277$. Also, we calculate τ-values based on $\hat{\alpha}$ values as $\hat{\tau}_1 = 0.2328296$ and $\hat{\tau}_2 = 0.2271088$. As seen in Table 4, Frank copula has the best fit (p-value=0.3594) for heart transplant data, because its p-value is higher than the other models. Gumbel copula model also cannot be rejected at the 5% significance level, however its p-value is lower than the Frank copula model. Clayton model is rejected at 5% significant level. So, we can conclude that the model has symmetric dependence structure for transplant and post-transplant survival times. Figures 1A-1C are constructed using λ-functions of Clayton, Gumbel and Frank copula model, respectively. These figures show which copula models have the better fit for the given data, visually. These figures are constructed from theoretical and empirical data to see how their fits are. In Figures, the left part is for the empirical λ-function of the data, the middle one is for the theoretical λ-function and the right one is for both empirical and theoretical λ-functions. When Figures 1A-1C are examined, it can be concluded that the Frank copula model has again the better fit to the data.

TABLE 4: The GoF results for Archimedean Copula models for the Stanford heart transplant program.

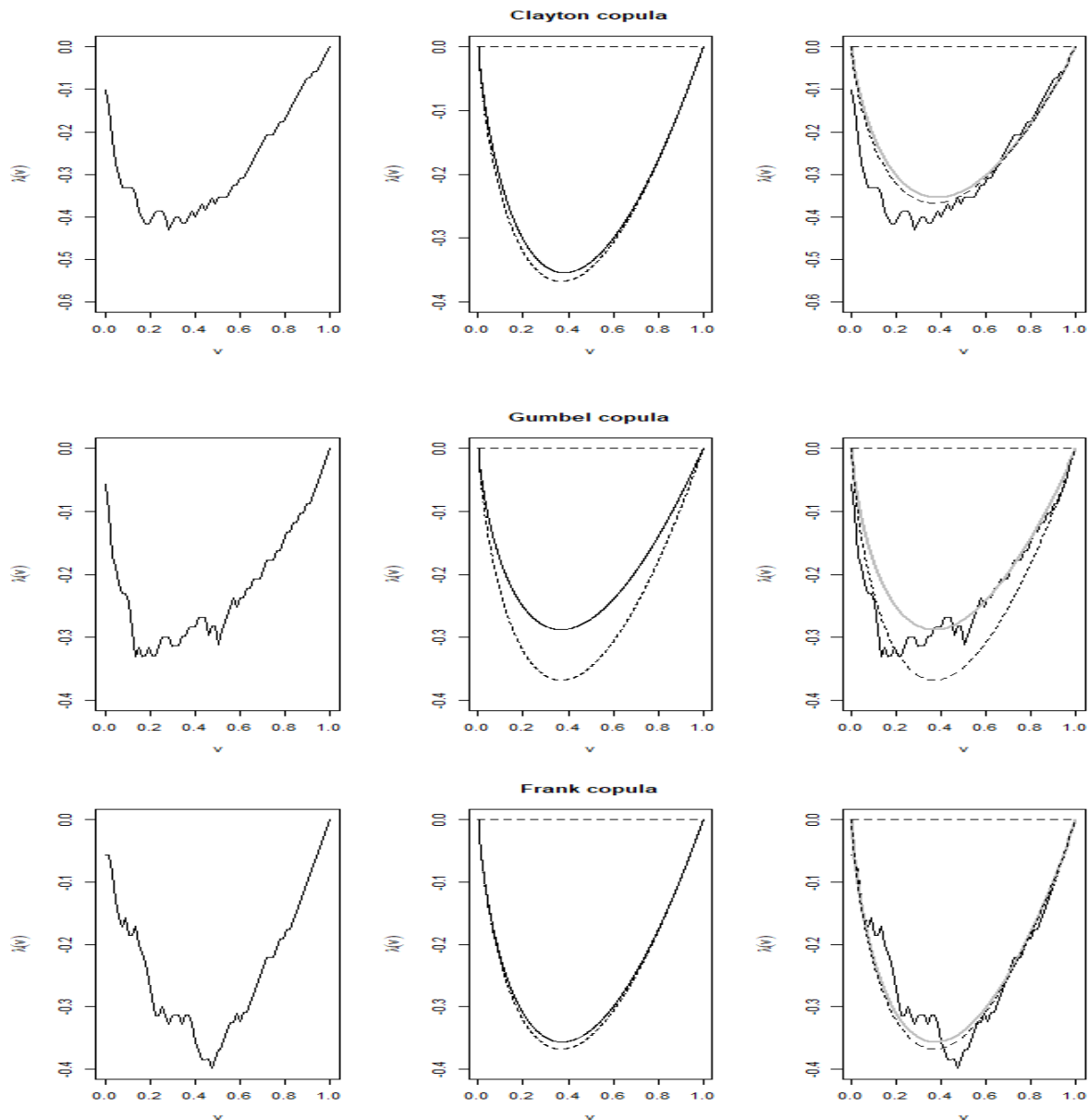|	$\hat{\alpha}_1$	$\hat{\alpha}_2$	$(\hat{\gamma}_1 - \hat{\gamma}_2)/\hat{\sigma}_{jack}$	p value
Clayton	0.1692321	0.6449004	-6.437196	0
Frank	2 .1930472	2.134277	0.3676287	0.3594
Gumbel	1.2789026	1.2598375	0.9582767	0.1711

FIGURE 1: λ-functions based on the Stanford heart transplant program for some copula models.

Secondly, our other data is obtained from diabetic retinopathy study.[13] We tackled 197 patients for our censored study. One of the eyes of the patients was randomly selected for photocoagulation treatment and an experimental treatment was applied to one eye, a standard treatment to the other eye, of each patient enrolled. Manatunga and Oakes[13] concentrated the effect of laser- photocoagulation therapy. The data consists of 197 patients. 38 patients have failure in both eyes, 79 patients have failure in one eye and 80 of them has no failure. Considering censored status, we use the variables as treated and untreated eyes and also we deal with time to loss of vision or last follow-up of the patients as survival time. Descriptive statistics of diabetic retinopathy data is given in Table 5.

TABLE 5: Descriptive Statistics of Diabetic Retinopathy Data.

Diabetic retinopathy	Mean	Std	Var	Min	max	Median
Standard survival time	32.29	21.46031	460.545	0.3	74.93	32.63
Photocoagulation survival time	38.87	20.78826	432.1517	1.47	74.97	42.23

In this data, we model the dependence structure between the survival times of the two eyes of the same patient after laser- photocoagulation therapy to see how they affect each other. We again deal with Frank copula, Gumbel copula and Clayton copula models. We apply the goodness-of-fit procedure for these copula models. Table 6 shows the goodness-of-fit results. The weighted and unweighted dependence parameters, test statistics and their p-values are shown in Table 6. When we examine Table 6, Frank copula model has the better fit to our data (p-value=0.435 (Frank)), because its p-value is higher than the other models. Also, for the Frank copula model, we obtain $\hat{\alpha}_1 = 1.5647446$ and $\hat{\alpha}_2 = 1.6743183$. Clayton copula model also cannot be rejected at the 5% significance level, however its p-value is lower than Frank copula model. Gumbel model is rejected at 5% significance level. So, we can conclude that the model has symmetric dependence structure between the survival times of the two eyes of the same patient after laser-photocoagulation therapy. The visual comparisons are also given in Figures 2A-2C via λ-functions. These figures are constructed for theoretical and empirical data to see how their fits are. In Figures, the left part is for the empirical λ-function of the data, the middle one is for the theoretical λ-function and the right one is for both empirical and theoretical λ-functions. So, it can be concluded that the Frank copula model has again the better fit to the data.

Frank copula is important because it allows negative dependence between the marginals and also the dependence is symmetric in tails. In theory, Frank copula can be applied to the model outcomes with strong positive or negative dependence.[14]

TABLE 6: The GoF results for Archimedean Copula models for the diabetic retinopathy study.

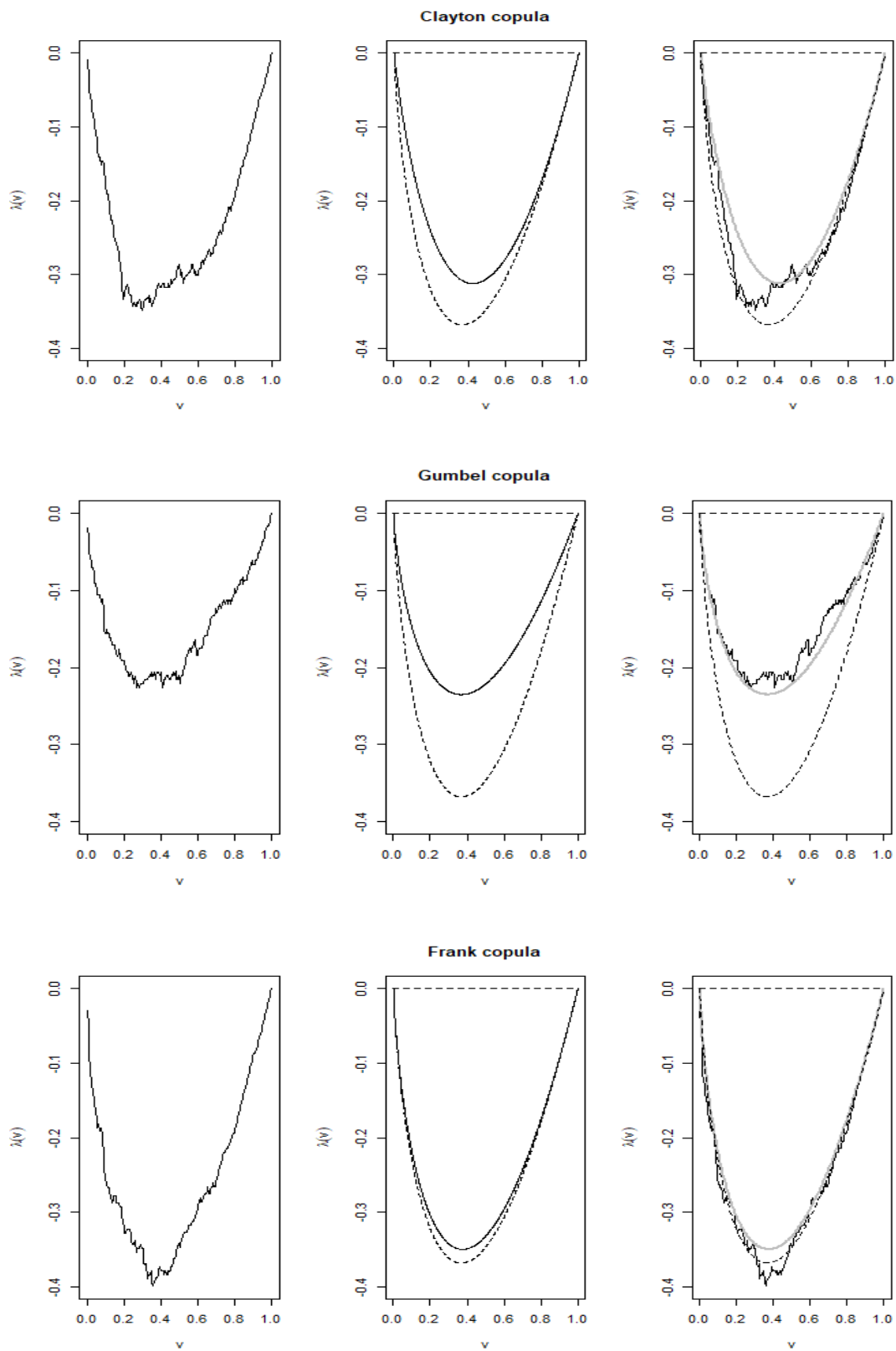|	$\hat{\alpha}_1$	$\hat{\alpha}_2$	$(\hat{\gamma}_1 - \hat{\gamma}_2)/\hat{\sigma}_{jack}$	*p value*
Clayton	1.164139	1.246918	-1.00857	0.1587
Frank	3.897294	3.944172	-0.1755508	0.435
Gumbel	1.5647446	1.6743183	-2.603582	0.0047

FIGURE 2: λ-functions based on the diabetic retinopathy study for some copula models.

Kendall distributions of the selected copula models for both of the data sets are visualized in Figure 3 and Figure 4. Kendall distribution plots show the distribution of the pairs of the data using Kendall function. These figures are plotted to see the goodness-of-fit results of the selected copula models to the data sets. It can also be concluded that Frank copula model has the better fit to both Stanford heart transplant data and diabetic retinopathy data.
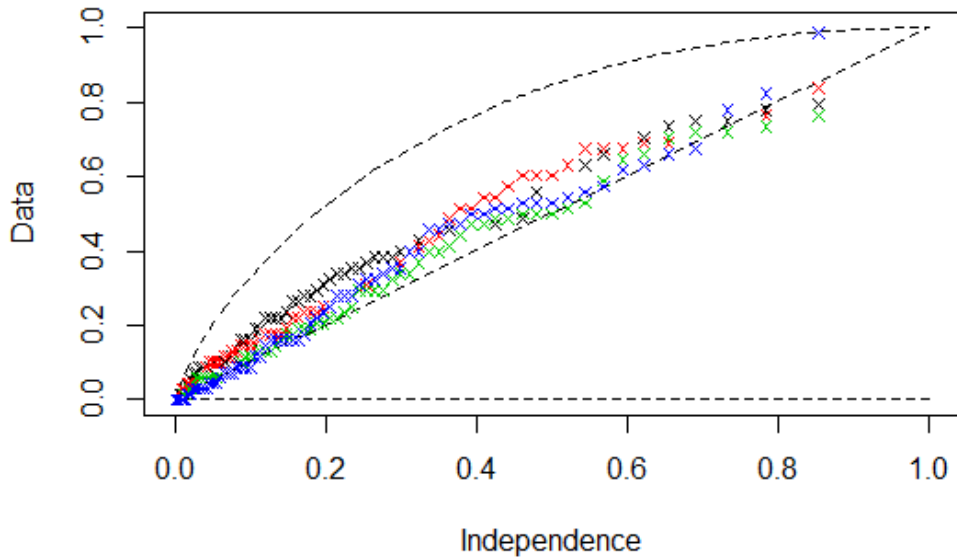
FIGURE 3: Kendall plot based on the Stanford heart transplant program
NOTE: Green (Frank copula), red (Clayton copula), blue (Gumbel copula)
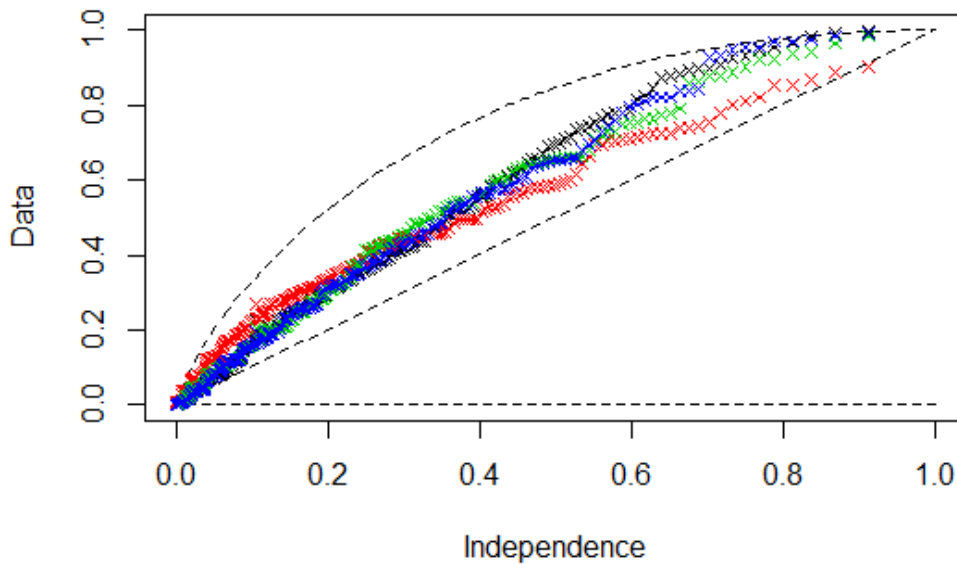

FIGURE 4: Kendall plot based on the diabetic retinopathy study
NOTE: Green (Frank copula), red (Clayton copula), blue (Gumbel copula)

Also, we calculate the Mean Squared Errors (MSEs) of the models for the data presented in Table 7. We can also conclude that Frank copula model also has minimum MSE values.

TABLE 7. MSE values for Heart transplant and Diabetic Retinopathy data.

MSE	Heart Transplant data	Diabetic Retinopathy data
Frank	0.0001510953	0.0005226535
Gumbel	0.0005936182	0.0005410127
Clayton	0.00097334402	0.00102377

The data pairs which exceed the critical layers are also obtained using Survival Hazard scenario which provides valuable resulting for assessing the probability of threatening events. After the model selection, survival Hazard Scenario is applied for heart transplant study and diabetic retinopathy study to determine the critical layers using the copula models. In Table 8, survival time after transplant and waiting time to transplant and its survival Hazard scenario levels α_t^k are given. Similarly, In Table 9, for non-treated (experimental) survival time and treated (photocoagulation) survival time of the pair of the eyes and their Survival Hazard Scenario levels α_t^k are given. The bivariate critical threshold for the Frank copula for given (x, y) is presented in the third column of Table 8 and Table 9. Also, α_t^k which shows the exceeding the critical layer probabilities are given in the fourth column. We can conclude that the exceeding probability of the critical layer 0.1527 is 0.3913 for the value of $(x, y) = (5,5)$.

TABLE 8: The probabilities that exceed the critical layer for the selected heart transplant data.

Survival time	Waiting time	$t = H(x, y)$	α_t^k
5	5	0.1527	0.3913
16	2	0.5326	0.9130
16	1	0.7826	0.9710
17	5	0.3804	0.7246
30	5	0.4750	0.8405
39	38	0.1156	0.3043
39	36	0.1386	0.3478
43	20	0.3190	0.6521
45	1	0.8382	0.9855
51	12	0.4527	0.7971

TABLE 9: The probabilities that exceed the critical layer for selected diabetic retinopathy data.

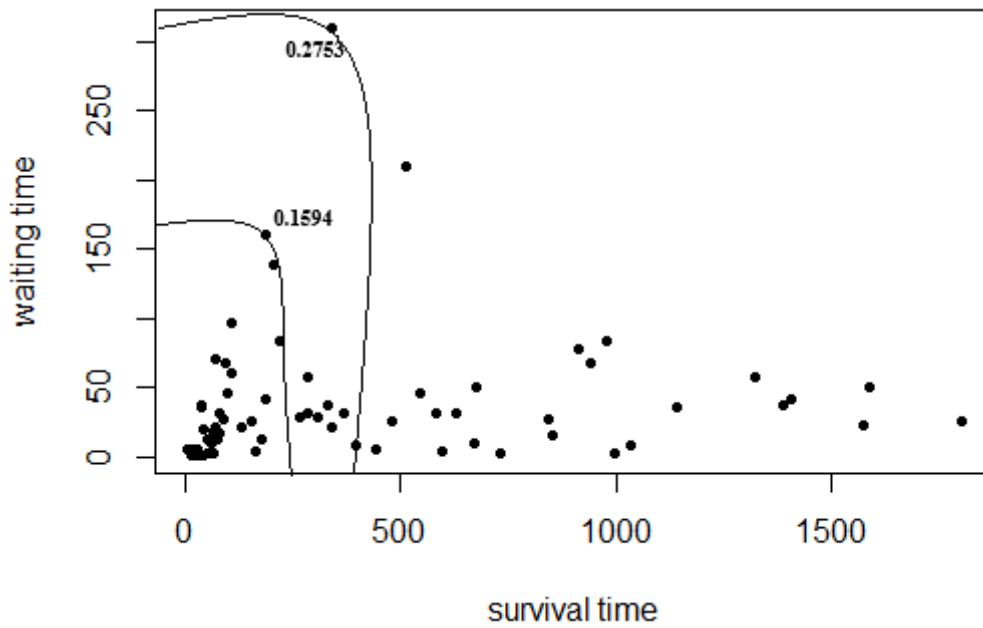| Non-treated Eye Survival Time	Treated Eye Survival Time	$t = H(x, y)$	α_t^k
46.23	46.23	0.2330	0.4771
31.3	42.5	0.2769	0.5431
42.27	42.27	0.2849	0.5532
20.6	20.6	0.4995	0.8172
0.3	38.77	0.3001	0.5736
54.27	65.23	0.0294	0.0152
10.8	63.5	0.0487	0.0913
23.17	23.17	0.4792	0.7868
13.83	58.07	0.0993	0.2385
48.53	46.43	0.2123	0.4365

FIGURE 5: Observed pairs of heart transplant data, selected critical layer L_t^{k} and level α_t^{k}.
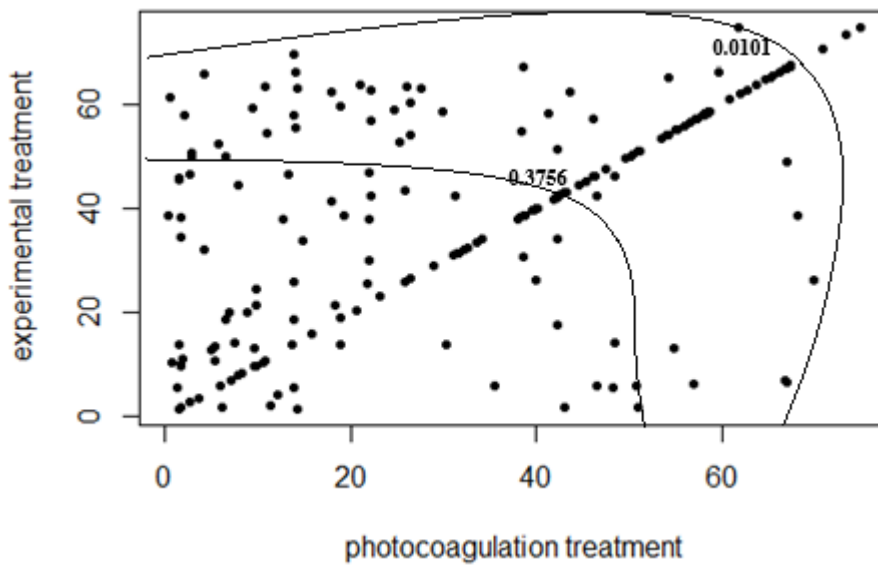

FIGURE 6: Observed pairs of diabetic retinopathy data, selected critical layer L_t^{k} and α_t^{k}.

In Figure 5 and Figure 6 we can see the observed pairs of the data, the selected critical layer L_t^{k} and the level α_t^{k}. From the figures, we can determine the critical layers, also we can determine which observations exceed the critical layers. For example, when survival time is 340 days and waiting time is 310 days, the risk level is 27.5% for heart transplant study. Similarly, in diabetic retinopathy study, when non-treatment time is 66.2 months and treatment time is 66.2 months, the risk level is 1%.

CONCLUSION

In this study, we deal with modeling and analyzing bivariate survival uncensored and right-censored data for three Archimedean copula models. We apply the goodness-of-fit method proposed by Emura et al.[1] to select the best appropriate Archimedean copula model to our data sets. Different from the conventional methods in survival analysis, this method provided a formal GoF test using weight function based on conditional likelihood. This strategy was also used by Oakes[2].

We use Kendall distribution function and λ-functions to choose the best fit visually. Since the variance estimator is complicated for the right-censored data, the jack-knife estimates are obtained. For both of the heart transplant data and diabetic retinopathy data, Frank copula model is selected. This means that the data sets have symmetric dependence structure. For the heart-transplant data, we investigate the relation between the waiting time for transplant and post-transplant survival time. As discussion in the paper Aitkin et al.[15], they concluded that the transplant appears to prolong survival. However, a selection bias in the procedure has been appeared. The non-transplant patients are those who die while waiting for a heart, while transplanted patients are those who survive until a suitable heart is found. So, from the dependence structure, we can conclude that when the waiting time for transplant gets longer and the patient is still alive, this means that the patient has relatively longer survival time after transplant, because the patient has low risk (stronger). So, very high risk patients generally do not survive the waiting period and so this confirms our results.

Also, in the diabetic retinopathy study given by Manatunga and Oakes[13], one of the eye of the patients was selected for laser photocoagulation treatment and the other eye was observed without treatment. As censored status, we use the variables as treated and untreated eyes. We model the dependence structure between the survival times of the two eyes of the same patient after laser-photocoagulation therapy, so we obtain the copula of related survival times. After modelling the dependence of variables we can conclude that both of the eyes have similar survival times after the laser- photocoagulation therapy. One of the eyes survive as the other one survives after the treatment.

Finally, a Hazard Scenario approach is also applied in evaluation the joint risk of the variables. We calculate the probability that patients of the survival time and waiting time jointly exceed the critical layer in heart transplant program and also that patients of the treatment survival time and non-treatment survival time jointly exceed the critical layer in diabetic retinopathy program. We deal with the Survival Hazard scenario to determine appropriate probability levels of some extreme events for both of the data. Determining the critical layers help the researcher to see which pairs are in extreme area or which pairs are in safe area.

Conflict of Interest

No conflicts of interest between the authors and / or family members of the scientific and medical committee members or members of the potential conflicts of interest, counseling, expertise, working conditions, share holding and similar situations in any firm.

Authorship Contributions

Idea/Concept: Burcu Hüdaverdi; Design: Burcu Hüdaverdi; Control/Supervision: Burcu Hüdaverdi; Data Collection and/or Processing: Ece Görceğiz; Analysis and/or Interpretation: Ece Görceğiz, Burcu Hüdaverdi; Literature Review: Ece Görceğiz, Burcu Hüdaverdi; Writing the Article: Ece Görceğiz, Burcu Hüdaverdi; Critical Review: Burcu Hüdaverdi; References and Fundings: Ece Görceğiz; Materials: Ece Görceğiz, Burcu Hüdaverdi.

▌REFERENCES

1. Emura T, Lin C, Wang W. A goodness-of-fit test for Archimedean copula models in the presence of right censoring. Computational Statistics & Data Analysis. 2010;54(12):3033-43. [Crossref]

2. Wang W, Wells M. Model selection and semiparametric inference for bivariate failure-time data. J Am Stat Assoc. 2000;95(449):62-72. [Crossref]

3. Nelsen RB. An introduction to copulas. 2nd ed. New York. USA: Springer; 2006. p.272.

4. Oakes D. A model for association in bivariate survival data. J Royal Stat Soc. Series B (Methodological). 1982:44(3);414422. [Crossref]

5. Emura T, Wang W, Hung HN. Semi-parametric inference for copula models for truncated data. Stat Sin. 2011;21:349-67.

6. Genest C, Rivest LP. Statistical inference procedures for bivariate Archimedean copulas. J Am Stat Assoc. 1993;88(423):1034-43. [Crossref]

7. Shih JH. A goodness-of-fit test for association in a bivariate survival model. Biometrika. 1998;85(1):189-200. [Crossref]

8. Clayton DG. A model for association in bivariate life tables and its application to epidemiological studies of familial tendency in chronic disease incidence. Biometrika. 1978;65(1):141-51. [Crossref]

9. Salvadori, G, Durante F, De Michele C. Multivariate return period calculation via survival functions. Water Resour Res. 2013;49(4):2308-11. [Crossref]

10. Salvadori G, Durante F, De Michele C, Bernardi M, Petrella L. A multivariate copula-based framework for dealing with hazard scenarios and failure probabilities. Water Resour Res. 2016;52:3701-21. [Crossref]

11. Susam SO, Ucer Hudaverdi B. Modeling the dependence structure of CO_2 emissions and energy consumption based on the Archimedean copula approach: the case of the United States. Energy Sources, Part B: Economics, Planning, and Policy. 2019;14(6):274-89. [Crossref]

12. Crowley J, Hu M. Covariance analysis of heart transplant survival data. J Am Stat Assoc. 1977;72(357):27-36. [Crossref]

13. Manatunga AK, Oakes D. Parametric analysis of matched pair survival data. Lifetime Data Analysis. 1999;5:371-87. [Crossref] [PubMed]

14. Trivedi PK, Zimmer DM. Copula modeling: an introduction for practitioners. Found Trends Economet. 2005;1(1):1-111. [Crossref]

15. Aitkin M, Laird N, Francis B. A reanalysis of the stanford heart transplant data. J Am Stat Assoc. 1983;78(382):264-74. [Crossref]