Brazilian Journal of Biometrics, **42**, 50-58 (2024)
doi:10.28951/bjb.v42i1.644

ARTICLE

A bivariate survival model for events with dependent failure times based on Archimedean copula functions. Application case: A sample of HIV patients

Jesús Alberto Peña-Guillén[1], Josefa Ramoni-Perazzi*[2] and Giampaolo Orlandoni-Merli[3]

1 Universidad de Los Andes, Mérida, Venezuela
2 Universidad Industrial de Santander, Bucaramanga, Colombia
3 Universidad de Santander, Bucaramanga, Colombia
*Corresponding author. Email: jramonip@uis.edu.co

(Received: November 17, 2022; Revised: June 3, 2023; Accepted: July 17, 2023; Published: March 15, 2024)

Abstract
This paper proposes a bivariate survival model for dependent failure times based on copula functions of the Archimedean family and the mean cumulative function for non-recurrent events of different types (MCFR̄E) and uses it to estimate the probability of survival from the occurrence of events of different types on the same HIV/AIDS patient. The copula functions evaluate the dependence structure between the failure times of the events experienced by the same patient throughout their follow-up period, and the MCFR̄E generates the marginal survival function for each event. The marginal function is a nonparametric estimator that gives the same estimated survival probability as the Kaplan-Meier estimator if the failure times of the different types of events are independent. If each patient experiences at least one event, a subset of them generates a compound event that affects the estimated probability of survival. The results show that the traditionally estimated survival probabilities are biased if dependent failure times are treated as independent.

Keywords: Archimedean copula family; Bivariate survival model; Dependent failure times; HIV/AIDS

1. Introduction

The applications of statistical survival models to data analysis in the clinical area have increased considerably in recent years. These consist of observing the moment or instants of time in which each patient experiences one or more events simultaneously. In these cases, classical models such as Kaplan & Meier (1958) or competing risk analysis are applied (Pintilie, 2006). If the occurrence of several events in the same patient is analyzed, these models do not capture the true survival probability because the time of occurrence of the different events depends on each other. This is an important characteristic that is ignored in those models that yield biased results. When considering the occurrence of several events on the same patient, the assumption of independence between the failure times of each event must be carefully revised since estimates

can be biased if the dependence between them is ignored (Taylor & Peña, 2013 and Boracchi & Orenti, 2015). When considering different types of events experienced by the same patient during his/her follow-up period, a partial history of the event is observed, so that independence between failure times of different events can rarely be assumed (Boracchi & Orenti, 2015).

Unlike the competing-risk events considered by Zheng & Klein (1995) and Lo & Wilke (2010), we propose an alternative survival model for different events that generate another one whose failure times are not independent based on copula functions of the Archimedean family and the history of different events experienced by the same statistical unit, defined here as the mean cumulative function for non-recurrent events of different types (MCFR$\overline{\text{E}}$). On the other hand, a new event can arise from the occurrence of a sequence of non-terminal events of different types on the same statistical unit, which we define as a compound event. This compound event is the result of the accumulation of individual events and has a different effect than that observed for each event. Its structure resembles a series system, a parallel structure system, or a system of the r-out-of-k form. Based on the characteristics of the available dataset, the compound event is studied here as a series system. If failure times depend on each other, the survival probabilities estimated by the proposed model are expected to differ from those obtained by classical survival models; otherwise, the estimates coincide. To our knowledge, there are no papers considering dependent failure times based on copula functions of the Archimedean family and the mean cumulative function for non-recurrent events of different types.

The proposed model is applied to a cohort of N=115 patients with a positive diagnosis of the human immunodeficiency virus (HIV), suffering from hepatitis B and under clinical treatment. They all attended the *Hospital Universitario de Los Andes* in Mérida-Venezuela every six months for their respective clinical check-ups during the follow-up period from January 2007 to December 2013. These patients were medicated with antiretrovirals according to the criteria established by the national health authority following the World Health Organization (WHO, 2013) for this type of mandatory reporting of infections (see Timaure, 2017).

The objective is to estimate their probability of survival from the occurrence of at least one of the two events related to the clinical manifestations of HIV, which in the long run may generate the acquired immunodeficiency syndrome (AIDS) event. For each patient, the events occurred at different instants of time during the follow-up period. The survival probabilities estimated with the proposed model are compared with the results of the Kaplan- Meier model. Our results show that if dependent failure times are treated as independent, the traditional estimates are biased.

2. Methodology: Description of the Proposed Survival Model

2.1 The Compound Event

In the clinical area, some pathologies depend on the occurrence of several events experienced by the same patient over time. This result is considered a compound event. Suppose that each patient i=1, 2, …, N is exposed to experience k different events, k=1, 2, …, K, at different instants of time, say T_{i1}, T_{i2}, …, T_{ik}, with a cumulative distribution function $F_k(t_k)$ and a survival function $S_k(t_k)$.

The compound event consists of a subset of events of different types k that act simultaneously on the same patient and has the property of producing another different effect. Each patient can experience a maximum of K_i events, with $K_i \leq K$; the compound event will be made up of r individual events of order k, with $t \leq K_i$. If r=k, the compound event consists of all k different events. If r=1, the compound event consists of at least one event; that is, there must be at least one event k for the compound event to occur. If r=g<K, the compound event consists of a subset of at least g events of different types k.

The proposed model and the estimated survival probability vary according to the composition of the compound event. Following Peña *et al.* (2018), this event depends on the objective of the

study, the meaning of the operational definition, and the area to which it is applied.

2.2 The Compound Event as a Serial System

Suppose that the *i*-th patient is exposed to the occurrence of *k* different events, but it is enough for at least one of them (*r*=1) to conclude that the composite event has occurred or been diagnosed during the follow-up period [0, δ_i), where δ_i is the maximum time of observation for the *i*-th patient. The patient survives if he/she recovers from each of the *k* different events. Let F_T^S be the probability that the compound event occurs under a series system, $F_T^S = \mathrm{P}[\bigcup_{k=1}^{K} T_k \leq t_k]$ which implies that the probability of surviving the occurrence of the k events is given by

$$\boldsymbol{S_{\cap T}(t) = \mathbf{P}[T > t] = \mathbf{P}\left[\cap_{k=1}^{K} T_k > t_k\right].} \quad (1)$$

If the terms T_k are mutually independent, then $S_{\cap T}(t) = \mathrm{P}[T > t] = \prod_{k=1}^{K} S_k(t_k)$. If T=min{T_1,T_2,...,T_k}, then $S_T(t) = \mathrm{P}[T > t] = \prod_{k=1}^{K} S_k(t)$. If the random variables T_1, …, T_k are identically distributed, then $S_k(t)$=$S(t)$, $\forall k$, so that $S_T(t) = [S(t)]^k$. However, if the terms T_k are mutually dependent, that is, a dependence structure between the instants in which the different events occur is described, then the following cases can be considered based on the copula functions (Nelsen, 2006) and from (1):

a) For *r*=*K*=2, the copula survival function is

$$S_{\cap T}(t) = \sum_{k=1}^{K} S_k(t_k) - (K-1) + \mathrm{C}_\theta\left(\cap_{k=1}^{K}\left(1 - S_k(t_k)\right)\right), \quad (2)$$

where C_θ is the copula function that describes the degree of dependence among the failure times T_k.

b) For *r*=*K*>2, it is not a simple generalization of (2) based on the properties of copula functions explained in Nelsen (2006), Úbeda (2001), and Peña (2018).

2.3 The MCFR̄E Function

The MCFR̄E function is the mean cumulative function of non-recurring events of different types experienced by the *i*-th patient during his/her follow-up period [0,δ_i). At each instant of time T_{ik}, the individual accumulates events and generates an accumulated "history" of events of different types (FAHD) whose function is given by $Y_i^*(t) \in \{0,1,\ldots,K_i\}$.

Let $y_{ik}^*(t)$ be the number of times each event occurs for the *i*-th patient during the instant of time *t*. Since they are non-recurring events of different types, each event will be observed only once, which implies that $y_{ik}^*(t)$=1, $\forall i,k$. Then, $\sum_{k=1}^{K_i} y_{ik}^*(t) = Y_i^*(t)$. Each function $y_i^*(t)$ is the "value" of each statistical unit; graphically, it takes the form of a "ladder", where each step is a different event e_k occurred at most at time δ_i.

For a population of failure times of N patients T_{ik}, the FAHD for each patient could include all the *k* different event types, a subset *r* of events of different types with $r < k$, or just one event type, which would lead to the conclusion that the *i*-th patient under observation has experienced the compound event. The structure of the MCFR̄E function looks like the MCF function, but the latter is for recurrent events of the same type, as discussed in Nelson (2003). The graphical representation of the FAHD function for the *i*-th individual is shown in Figure 1.

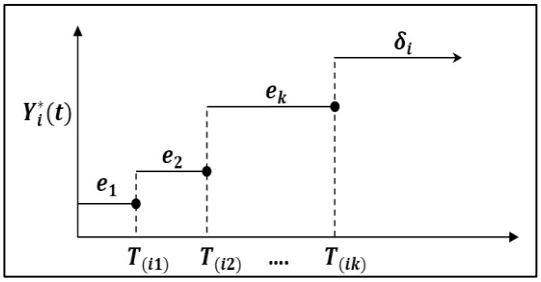

Figure 1. Cumulative history function for events of different types experienced by the i-th patient.

Since it is a serial system, the occurrence of at least one event of the form e_k during the follow-up period of the *i*-th patient is enough to conclude that the compound event has occurred. Let $M^*(t)$ be the MCFR$\bar{E}$ function defined as the mean cumulative number of events of different types experienced by each patient, thus $M^*(t) = \frac{1}{N}\left(Y_1^*(t), Y_2^*(t), \ldots, Y_N^*(t)\right)$. As in the MCF function, the statistical model for a population of *N* units is the FAHD (Peña, 2018). Thus, the model is represented by a population of *N* curves (see Figure 2).

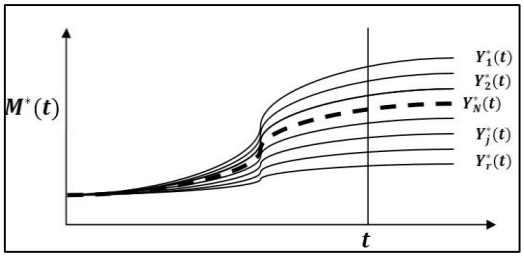

Figure 2. Mean cumulative function of non-recurring events of different types (MCFR$\bar{E}$).

Based on Figure 2, at each instant of time $t_l=t$ the values of each FAHD may have different probability distributions with mean $M^*(t)$ (dashed line). As for the MCF function for a very large population of values, this distribution is continuous with the same characteristics of a continuous density function in the sense that $Y^*(t)$, understood as a stochastic process, converges towards a Gaussian process according to the Central Limit Theorem. Thus, the derivative $\partial(\partial t\, M^*(t)=m^*(t))$ is assumed to exist and represents the instantaneous population rate of occurrence of events where at least one is of a different type because it depends on the instant of time in which an event is experienced. The function $m^*(t)$ and the "failure rate" $\lambda(t)$ related to the distribution of the lifetime of a patient for a terminal event are different. In this framework, $m^*(t)$ measures the risk of experiencing different non-terminal events, and $\lambda(t)$ is the risk of failure function inversely related to the survival function for a given time. Due to the characteristic of non-terminal events, $m^*(t)$ is a function that allows the patient to be observed several times during the follow-up period $[0,\delta_i)$, while for the function $\lambda(t)$ the patient is observed only once when the event of interest occurs.

The *i*-th patient is exposed to experiencing K_i different events at most, such that *G* is the subset of events that constitute the compound event $G=\{1, 2, \ldots, K_i\}$. If $G= 1$, the occurrence of at least one out of *G* possible events a patient can experience generates a compound event; thus, there are $\binom{K_i}{1}$ possibilities of observing this compound event. If $G= 2$, the occurrence of at least two events is the compound event. In general, if $G=g$, the occurrence of at least *g* events generates the compound event, but if $G=K_i$ the occurrence of all the different events generates the compound event. Consequently, each subset *g* of events is an MCFR$\bar{E}$ function, denoted by $M_g^*(t)$. At time *t* for the *i*-th patient, there is a mean cumulative number of events of different types for the *N* population values of the form $Y_{ig}^*(t)$, which implies that $M_g^*(t)= E_i\{Y_{ig}^*(t)\}$, being the classical arithmetic mean $\widehat{M}_g^*(t) = \bar{Y}_{ig}^*(t) = \frac{1}{N}\sum_{i=1}^{N} Y_{ig}^*(t)$ an estimator of $M_g^*(t)$. To estimate $M_g^*(t)$ in each *t*, the times where the occurrence of an event *k* was observed are ordered in increasing magnitude, along with the instants of time where the occurrence of some event was not observed, that is, the censored times. Assume there are *L* failure times, $t_0=0 <t_1<\cdots<t_l<\cdots<t_L$. Then, in each $t(l)$ there exists a given number of patients at risk of experiencing a new event of a different type, denoted as $Y^*(t(l))$. If $t(l)$ is censored, then $Y^*(t(l))=N-1$; otherwise, $Y^*(t(l))=N$. For each instant of time $t(l)$, $Y_k^*(t_{(l)})$ denotes the number of times a type *k* event occurred, which implies that $m_k^*(t) = Y_k^*(t)/Y^*(t)$, where $m_k^*(t)$ is the average rate of patients who experienced the event *k* at the instant of time $t(l)$. Therefore, the sampled value of $M_k^*(t)$ at each $t(l)$ is obtained by adding the preceding increments of $m_k^*(t)$, $\widehat{M}_k^*(t) = \sum_{\{t_{(l)}\le t\}} m_k^*(t)$.

If the population of patients is exposed to the occurrence of *K* events of different types, then the population has a mean *K*-variant curve given by $\left[\widehat{M}_1^*(t), \widehat{M}_2^*(t), \ldots, \widehat{M}_k^*(t), \ldots, \widehat{M}_K^*(t)\right]$. If the

compound event G is made up of g events of different types, then $\widehat{M}_G^*(t)$ with $G \leq K$ consists of the sum of these functions of each event g, in other words, $\widehat{M}_G^*(t)=\sum_{g\in G} M_g^*(t)$. Peña *et al.* (2018) show that the marginal survival probability for each subset g of k events is $S_T(t) = \exp\left(-M_g^*(t)\right) = \exp(-\sum_{k\in G} M_k^*(t))$. Furthermore, if the failure times for each event k are independent, $S_T(t)$ coincides with the nonparametric Kaplan-Meier estimator, $S_{KM}(t)$. Therefore, the marginal survival functions for the different events k are given by

$$S_k(t_k) = \exp\left(-M_k^*(t_k)\right). \tag{3}$$

2.4 Proposed Model from the Copula Functions of the Archimedean Family

To each set of failure times T_k, corresponds a marginal survival function $S_k(t)$ given by (4), introduced next. Although each T_k can correspond to a parametric distribution function, for example, exponential or Weibull, the joint fit between T_k and T_j with $k\neq j$ does not necessarily remain the same. According to the functional form of the dependence structure between T_k and T_j, not necessarily linear, a particular copula will adjust. The random behavior of the marginal function does not imply that it is the same for the multivariate survival function, hence the use of the non-parametric survival function (4) as the marginal function of the proposed model. If the failure times of events of different types are associated, one way to deal with the problem is through copula functions, since they have the property of connecting marginal survival functions. The degree of association between failure times, θ, is measured through Kendall's *tau* coefficient as a measure of the dependence among copula functions (Nelsen, 2006).

For K=G=2, each patient is exposed to the occurrence of two different events, but the occurrence of the compound event depends on each patient experiencing at least one. Let T_{ik} with i=1, 2, …, N and k=1, 2, be the failure times of both events for the i-th patient. According to Nelsen (2006) and Genest & Mackay (1986), if $(U_1, U_2)'$ is a bivariate random variable with standard uniform marginals [0,1], the copula function is defined as $\mathrm{C}(u_1, u_2) = \mathrm{P}[U_1 \leq u_1, U_2 \leq u_2]$.

The copula function that fits the dependence structure between the random variables T_1 and T_2 cannot be chosen arbitrarily. In this paper, we only consider the cases where the "best" fit is given by copula functions of the Archimedean family with a strict generator and that share certain characteristics with the survival functions as a function of the dependence structure described between the T_k terms. According to Nelsen (2006), for a continuous, decreasing, and convex function $\phi(u,\theta)$ such that ϕ:(0,1]→[0,∞) with $\phi(1,\theta)$=0 as a strict generator, the Archimedean function copula is written as $\mathrm{C}_\theta(u_1, u_2) = \phi_\theta^{-1}\left(\phi(u_1, \theta) + \phi(u_2, \theta)\right)$.

Let T_1 and T_2 be non-negative random variables with survival functions $S_1(t_1)$ and $S_2(t_2)$, then U_1=$S_1(T_1)$ and U_2=$S_2(T_2)$. By the Integral Transformation Theorem, it follows that U_k~$U(0,1)$, with k=1,2, which defines the bivariate survival distribution function of $(U_1, U_2)'$ through a copula function to obtain the bivariate joint survival function of $(T_1, T_2)'$, that is, the proposed model. The different ways of expressing the proposed survival model using the Archimedean copula functions with a strict generator (Peña, 2018) based on the dependence structure described between the random variables with uniform distribution [0,1] are shown next:

Bivariate survival model based on the Clayton copula:

$$S_{12}(t_1, t_2) = \sum_{k=1}^{2} S_k(t_k) - 1 + \left\{\left(\sum_{k=1}^{2}\left(1 - S_k(t_k)\right)\right)^{-\theta} - 1\right\}^{-\frac{1}{\theta}}. \tag{4}$$

Bivariate survival model based on the Gumbel-Hougaard copula:

$$S_{12}(t_1, t_2) = \sum_{k=1}^{2} S_k(t_k) - 1 + exp\left[-\left(\sum_{k=1}^{2}\left(-ln\left(1 - S_k(t_k)\right)\right)^{\theta}\right)^{\frac{1}{\theta}}\right]. \tag{5}$$

Bivariate survival model based on the Frank copula:

$$S_{12}(t_1,t_2) = \sum_{k=1}^{2} S_k(t_k) - 1 - \frac{1}{\theta} ln\left[1 + \frac{\prod_{k=1}^{2}\left(e^{-\theta\left(1-S_k(t_k)\right)}\right)}{e^{-\theta}-1}\right]. \quad (6)$$

Bivariate survival model based on Joe copula:

$$S_{12}(t_1,t_2) = \sum_{k=1}^{2} S_k(t_k) - \left[\sum_{k=1}^{2} S_k^{\theta}(t_k) - \left(\prod_{k=1}^{2} S_k(t_k)\right)^{\theta}\right]^{\frac{1}{\theta}}. \quad (7)$$

where each marginal $S_k(t_k)$ in each model is given by (3).

3. Results and Discussion

We consider a cohort of 115 patients with positive diagnoses of HIV and hepatitis B infection, followed every six months from January 2007 to December 2013. For each patient, we recorded both personal and clinical information regarding the occurrence of one or more events of different types at the time of consultation, which could generate another event of greater clinical importance, such as AIDS. The objective was to measure a patient´s probability of survival based on the occurrence of two events of different types recorded at different moments during their follow-up period.

HIV is an infection that attacks and destroys the immune system, mainly lymphocytes TCD4+ and TCD8+ as the plasma viral load (PVL) increases. The biomarkers TCD4+ and CVP are widely used for the evaluation of the progression of HIV infection. This infection goes through different clinical stages, the most serious being AIDS, characterized by profound immunosuppression. In this study, the clinical structure of the composed event HIV was defined based on the occurrence of two events: TCD+4<200 cells/mm3 (here defined as event 1) and/or CVP>100,000 (here defined as event 2). The occurrence of these events does not have a particular order, but the presence of at least one of them generates the composite event of AIDS. The CVP for each patient was measured using the polymerase chain reaction (PCR) technique, and the count of TCD4+ lymphocytes was measured by flow cytometry (Timaure, 2017).

3.1 Description of the Failure Times for Both Events

The observation time in which one or no event occurred is equivalent to the age of the patient at the moment of the first consultation. The failure times of patients who did not experience both events (patients who survived the AIDS event) are censored to the right. If, at the time of consultation, the patient has experienced at least one of the events (1, 2, or both), his failure time is the time of occurrence of the AIDS event. For the statistical analysis, we first determine the univariate probability distribution that better fits the failure times of each event before fitting the copula function. In this case, the best fit for the occurrence times of each event is the Gamma probability distribution with estimated parameters (0.39, 16.18) and (0.45, 20.68), which points to the Gumbel-Hougaard copula as the best fit of the copula function. Figure 3 shows the positive dependence structure between both instants of time registered for each patient. The red dots on the scatter plot represent the occurrence of the compound event AIDS, while the blue dots indicate the censored events.

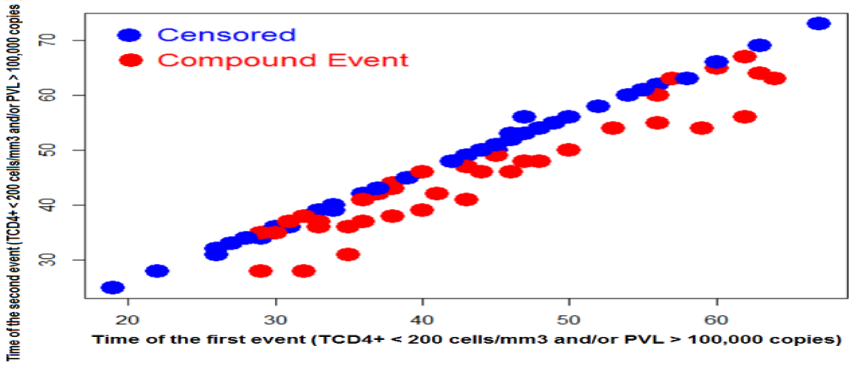

Figure 3. Dependence structure between failure times of events 1 and 2 in each patient.

When the patient experiences one event, the occurrence of the other is almost immediate, even though the dependence structure is not completely linear in both the first moments and the later ones. A kind of curve is observed at the points that indicate the occurrence of the composite event of AIDS (red dots). In this case, Kendall's *tau* correlation coefficient is equal to 0.83, indicating a strong relationship between both failure times.

3.2 Adjustment of the Copula Function for Failure Times of both Events and Estimation of the Probability of Bivariate Survival.

The BiCopSelect function from the VineCopula library of the R package version 4.0.3 suggests that the Gumbel-Hougaard copula function fits the best (p-value = 0.8937), which implies that the bivariate survival model for these data is given by (5) presented above (Genest *et al.*, 2013). Figure 4 shows the estimated bivariate survival curve identified by the proposed model (blue line). If the failure times of each type of event are independent, the Kaplan-Meier estimator is a good option to estimate the probability of survival. However, the probability of survival (red line) would be underestimated if the dependence between the two failure times was ignored. If the failure times were independent, both survival curves would coincide. As can be seen, the survival curve obtained from the proposed model is above the one obtained using the Kaplan-Meier model, showing the effect of the dependence between failure times. In addition, by considering the existing dependence between the failure times, a better survival prognosis is obtained for the HIV patients in the sample. For young patients under 30 years of age, the survival prognosis using both models is equivalent, but a better survival prognosis is obtained from the proposed model as the patient gets older; for patients older than 60 years, the probability of survival tends to stabilize above 40%. Table 1 summarizes some survival probabilities estimated by both models, given that the composite AIDS event occurred in 45% of the patients in our sample.

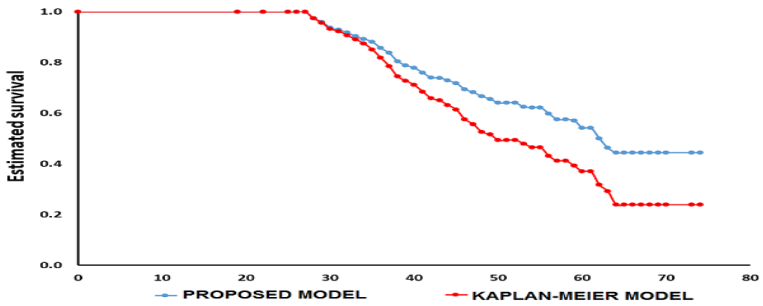

Figure 4. Estimated survival functions for dependent failure times (blue curve) versus independent failure times (red curve) for events 1 and 2.

Table 1. Estimated survival function using the Kaplan-Meier model (KM) and the proposed model (PM)

Model	Value of t_k					
	<25	35	45	55	65	75
KM	1.0000	0.8508	0.6133	0.4653	0.2390	0.2390
PM	1.0000	0.8809	0.7176	0.6225	0.4445	0.4445

Note: Estimated probabilities of survival based on a sample of 115 patients with a positive diagnosis of HIV and hepatitis B infection observed during the period 2007- 2013.

Based on these results, the estimated probability of survival of a 35-year-old patient is 0.8809, which larger compared to the one estimated by the Kaplan-Meier model. The older the patient, the lower the probability of survival, always higher in our model; after 65 years of age, it tends to stabilize at 0.4445. These results agree with the structure of the data set, indicating the failure time for each individual corresponding to each type of event, as well as the censored data for each event and the composite event.

4. Conclusions

This paper uses a bivariate survival model based on the copula functions of the Archimedean family to analyze the occurrence of two events of different types in the same patient, considering the possible dependence between failure times. Copula functions can capture the true dependency structure between two or more continuous random variables. Dependence structures with a positive, negative, or no trend are analyzed, which are the usual functional forms shown by the failure times of two events of different types, and represent how each patient experiences at least one event of a different type. These structures depict how each patient experiences at least one event of a different type. Either the greatest dependence may occur at the beginning during the first moments and considerably decrease in the last ones, or the greatest dependence may be observed during the last moments, or at the central failure times with a very low degree of dependence towards the extremes.

If the trend in the degree of dependence is positive and almost perfect, it can be concluded that the occurrence of one event produces almost immediately the occurrence of the other event; if it is negative and almost perfect, the early occurrence of one event delays the occurrence of the other. Hence the use of the copula functions of the Archimedean family to generate the proposed model. An important property of the proposed model is that, if failure times are independent, it is simplified to the copula function of the product, which coincides with the non-parametric Kaplan-Meier estimator.

If the patient experiences several non-recurrent events of different types during his/her follow-up period $[0, \delta_i)$, a subset of them may have the property of generating the composite event AIDS, whose occurrence depends on the presence of at least one of two possible conditions of this pathology. Thus, the proposed survival function yielded the best survival prognosis conditioned on the dependence structure of the events that produced the composite one. When compared with a classic model such as Kaplan-Meier, we determined that both sets of failure times considered in this case were dependent. If both the Kaplan-Meier survival curve and the proposed survival curve coincide, it is possible to conclude that the failure times of the different types of events are independent; otherwise, the curves would show different survival prognoses due to the dependence effect between the times of failure of events of different types. Therefore, through the proposed bivariate survival model, it is possible to capture the true dependence structure that exists between the failure times of the different types of non-recurring events that generate the composite event. Furthermore, when considering the occurrence of several events on the same individual during different instants of time, the best survival prognosis is not given by the highest probability of survival that a model can give us, but by the best structure of dependence that is captured between the failure times of the different types of events.

As for the selection of the copula function, which is part of the argument of the proposed bivariate survival function, it is based on several factors, such as the degree of dependence between the failure times of each type of event, the dependence structure described between them, and the marginal functions that are part of their structure. In this way, the probability distribution is obtained based on a particular copula function that finally constitutes the proposed model to analyze probabilities corresponding to dependent random variables.

Acknowledgments

In memorian of Professor Pedro Harmath who could not see the final product of our research.

References

1. Boracchi, P. & Orenti, A. Survival functions in the presence of several events and competing risks. Estimation and interpretation beyond Kaplan-Meier. *International Journal of Statistics in Medical Research* **4,** 121-139 (2015).

2. Genest, C., Huang, W. & Dufour, J. A regularized goodness-of-fit test for copulas. *Journal de la Société Francaise de Statistique* **154,** 64-77 (2013).
3. Genest, C. & Mackay, J. The joy of copulas: Bivariate distributions with uniform marginal. *The American Statistician* **40,** 280-283 (1986).
4. Kaplan, E. & Meier, P. Nonparametric estimation from incomplete observations. *Journal of the American Statistical Association* **53** (282), 457-481 (1958).
5. Lo, S. & Wilke, R.A. A copula model for dependent competing risks. *Journal of the Royal Statistical Society Series C Applied Statistics* **59** (2), 359 – 376 (2010).
6. Nelsen, R. *An introduction to Copulas* (Springer, 2006).
7. Nelson, W. *Recurrent Events Data Analysis for Product Repairs, Disease Recurrences, and Other Applications* (ASA-SIAM, 2003).
8. Peña, J. Modelos de supervivencia bivariantes para tiempos de fracaso posiblemente dependientes de eventos de distinto tipo experimentados por un mismo individuo (Doctoral Dissertation in Statistics, Universidad de Los Andes, Venezuela, 2018).
9. Peña, J., Ramoni-Perazzi, J. & Orlandoni-Merli, G. Un modelo de supervivencia bivariante para eventos dependientes bajo el enfoque de funciones cópulas. Revista de la Facultad de Farmacia 60, 13-24 (2018).
10. Pintilie, M. *Competing Risk: A Practical Perspective* (Wiley, 2006).
11. Taylor, L. & Peña, E. Parametric estimation in a recurrent competing risks model. *Journal of the Iranian Statistical Society* **20**, 514-537 (2013).
12. Timaure, R. *Método unificado para el estudio de la sensibilidad a los datos faltantes y las desviaciones a los supuestos de la distribución en modelos de efectos mixtos* (Doctoral Dissertation in Statistics, Universidad de Los Andes, Venezuela, 2017).
13. Úbeda, M. *Cópulas y Cuasicópulas: Interrelaciones y nuevas propiedades. Aplicaciones* (Doctoral Dissertation in Statistics, Universidad de Almería, España, 2001).
14. WHO. *Directrices unificadas sobre el uso de medicamentos antirretrovíricos para el tratamiento y la prevención de la infección por el VIH, Venezuela 2013.* Available at (http://www.who.int/hiv/pub/).
15. Zheng M, & Klein J.P. Estimates of Marginal Survival for Dependent Competing Risks Based on an Assumed Copula. *Biometrika* **82** (1), 127-132 (1995).