

Article

# Computation of the Mann–Whitney Effect under Parametric Survival Copula Models

Kosuke Nakazono <sup>1,2</sup>, Yu-Cheng Lin <sup>3</sup>, Gen-Yih Liao <sup>3</sup>, Ryuji Uozumi <sup>2</sup>  and Takeshi Emura <sup>1,4,\*</sup> 

<sup>1</sup> Research Center for Medical and Health Data Science, The Institute of Statistical Mathematics, Tokyo 190-8562, Japan; nakazono@ism.ac.jp

<sup>2</sup> Department of Industrial Engineering and Economics, Tokyo Institute of Technology, Tokyo 152-8552, Japan; uozumi.r.ab@m.titech.ac.jp

<sup>3</sup> Department of Information Management, Chang Gung University, Taoyuan 33302, Taiwan

<sup>4</sup> Biostatistics Center, Kurume University, Kurume 830-0011, Japan

\* Correspondence: takeshiemura@gmail.com

**Abstract:** The Mann–Whitney effect is a measure for comparing survival distributions between two groups. The Mann–Whitney effect is interpreted as the probability that a randomly selected subject in a group survives longer than a randomly selected subject in the other group. Under the independence assumption of two groups, the Mann–Whitney effect can be expressed as the traditional integral formula of survival functions. However, when the survival times in two groups are not independent of each other, the traditional formula of the Mann–Whitney effect has to be modified. In this article, we propose a copula-based approach to compute the Mann–Whitney effect with parametric survival models under dependence of two groups, which may arise in the potential outcome framework. In addition, we develop a Shiny web app that can implement the proposed method via simple commands. Through a simulation study, we show the correctness of the proposed calculator. We apply the proposed methods to two real datasets.

**Keywords:** censoring; copula; Hand’s paradox; Mann–Whitney effect; potential outcome; stress–strength model; survival analysis; survival function; treatment effect; two-sample comparison

**MSC:** 62-04; 62D20; 62F10; 62N01; 62N03; 62N05



**Citation:** Nakazono, K.; Lin, Y.-C.; Liao, G.-Y.; Uozumi, R.; Emura, T. Computation of the Mann–Whitney Effect under Parametric Survival Copula Models. *Mathematics* **2024**, *12*, 1453. <https://doi.org/10.3390/math12101453>

Academic Editor: Li-pang Chen

Received: 15 March 2024

Revised: 1 May 2024

Accepted: 6 May 2024

Published: 8 May 2024



**Copyright:** © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

When comparing the survival times of two independent groups, the Mann–Whitney parameter plays an important role in the two-sample problem [1]. The Mann–Whitney parameter, say,  $p$ , is defined as the probability that a random subject from one group (with survival time  $T_1$  in group 1) survives longer than an independent random subject from the other group (with survival time  $T_2$  in group 2), plus one-half the probability that the two subjects survive at the same time (tie):

$$p = \mathbb{P}(T_1 > T_2) + \frac{1}{2}\mathbb{P}(T_1 = T_2). \quad (1)$$

When researchers verify the superiority of a treatment over another, they examine whether the effect exceeds the null difference,  $p = 0.5$ . For instance, when they obtain  $p = 0.7$ , this effect is transformed to 0.8 of Cohen’s  $d$  according to well-known benchmark values [2]. Therefore, it can be interpreted as a large effect. The Mann–Whitney effect relates to important statistical ideas, such as, the Mann–Whitney test [3], hazard ratios, and win ratio [4]. The Mann–Whitney test examines the null hypothesis  $H_0 : p = 1/2$  vs.  $H_1 : p \neq 1/2$ . The hazard ratio is the main effect measure of the Cox proportional hazards model, which is a typical statistical model in survival analysis. The win ratio  $w$  is given by the odds of  $p$ ; that

is,  $w = p/(1 - p)$ . That is,  $p > 1/2$ , or equivalently  $w > 1$ , implies a protective survival effect for group 1.

The problem of estimating the parameter  $p$  plays an important part in survival and reliability analysis. The basic idea was first studied by Birnbaum [5]. They illustrated an attractive relationship between the Mann–Whitney statistic and the stress–strength model. Efron [1] first proposed a nonparametric estimator (Efron’s estimator) for  $p$  under independent censoring. Since then, this topic has been investigated by several researchers. In the following, we refer to some recent studies in the field of survival and reliability analyses. Dobler & Pauly [6] modified Efron’s estimator for  $p$  under small sample sizes. Emura & Hsu [7] proposed a copula-graphic estimator for  $p$  and suggested the Mann–Whitney test to compare two survival distributions in the presence of dependent censoring. Biswas et al. [8] introduced the Bayesian estimation of  $p$  for the log-Lindley distribution. Rubarth et al. [9] proposed estimating the Mann–Whitney effects in factorial clustered data. Hu et al. [10] developed methodologies for constructing fixed-accuracy confidence intervals of  $p$  when  $T_1$  and  $T_2$  follow geometric and the exponential distributions, respectively. de la Cruz et al. [11] studied a estimation procedure of the stress–strength model for the two independent unit-half-normal distributions with different shape parameters. Patil et al. [12] investigated the effect of dependence of the variables on  $p$  in the stress–strength model with the exponential margins. Nowak et al. [13] proposed a group sequential method for estimating the Mann–Whitney parameter. Singh et al. [14] studied the estimator of  $p$  in point, interval, and Bayesian estimations when the stress variables follow geometric and Lindley distribution. All the methods assumed that  $T_1$  and  $T_2$  are independent.

When  $T_1$  and  $T_2$  are independent of each other and continuous, one can estimate  $p$  with the marginal distributions based on the following integral:

$$p = - \int_0^{\infty} S_1(t) dS_2(t).$$

That is, one can estimate  $p$  by estimating two marginal survival functions,  $S_1$  and  $S_2$ . However, this is not the case when  $T_1$  and  $T_2$  are dependent; the phenomenon is sometimes called “Hand’s paradox” [15]. This showed that the paradox arises when  $T_1$  and  $T_2$  are regarded as potential outcomes in the causal inference framework. Therefore,  $p$  in the integral cannot be interpreted as the true treatment effect. Moreover, the dependence of outcomes from observation to observation is well known in factorial, paired, and cross-over designs [16,17].

Since  $p$  is not identifiable solely from independently sampled data, Fan & Park [18] suggested a bound for  $p$  under all possible dependence structures for  $T_1$  and  $T_2$ . Alternatively, Fay et al. [19] reformulated  $p$  such that it can be identified from randomized treatment assignments. Computing  $p$  under various dependence structures gives the researchers additional information for their decision making about the behavior of  $p$  when the outcomes are not independent. However, to estimate the true  $p$ , we must model the bivariate survival function of  $T_1$  and  $T_2$ . A copula is often used to model joint distributions of dependent survival times [20,21]. Here, the major challenge is to derive the formula of  $p$  under copula models and its extension to allow for a limited follow-up length for survival times.

In this article, we propose a model for the bivariate survival function by using parametric copulas and parametric marginal distributions. Our focus is to study  $p$  when there is dependence between  $T_1$  and  $T_2$ . Therefore, we focus on assessing how  $p$  vary under various dependences modeled by copulas. We then derive a new formula for computing  $p$  under a parametric copula. We also propose a new formula for  $p$  under the restricted follow-up. To make the proposed computation method for  $p$  to be easily performed by users, we develop a Shiny-based web app. Furthermore, we validate the accuracy of the proposed computation method and Shiny web app by simulations. We finally analyze real data, where the proposed copula-based estimators are compared with Efron’s estimator (benchmark).

The rest of the paper is organized as follows. In Section 2, we define copula-based models and introduce several well-known copula families. In this section, we show that one can compute  $p$  by a theorem given by Emura & Pan [22], and we extend the theorem to compute  $p$  when the follow-up time is restricted up to time  $\tau$ . In Section 3, we introduce a Shiny web app in the R software version 4.2.3 that can compute  $p$  via simple commands. In Section 4, we describe a simulation study to show the correctness of the proposed calculator for  $p$ . In Section 5, we illustrate a meaningful application of our proposed method using survival data. The computer code for the simulations and data analyses is available in Supplementary Materials.

## 2. Proposed Method

In this section, we first introduce copula-based bivariate survival models for  $T_1$  and  $T_2$ . We then propose our method for computing the Mann–Whitney effect  $p$  in Equation (1) under the copula models.

### 2.1. Survival Copula Models for Dependent Survival Time

According to Sklar [23], any bivariate distribution function for  $(T_1, T_2)$  can be formulated by using a copula. A bivariate copula is a bivariate distribution function for two uniformly distributed variables on  $[0, 1]$  [24,25]. In many applications of copulas on survival analysis, bivariate survival functions are modeled via survival copulas [20,24,26,27]. Below, we introduce a survival copula model for a bivariate survival function  $S(t_1, t_2) = \mathbb{P}(T_1 > t_1, T_2 > t_2)$ .

Let  $T_1$  and  $T_2$  be continuous survival times with marginal survival functions  $S_1$  and  $S_2$ , respectively. We model the bivariate survival function of  $T_1$  and  $T_2$  using a survival copula  $C$ :

$$\mathbb{P}(T_1 > t_1, T_2 > t_2) = C(S_1(t_1), S_2(t_2)). \tag{2}$$

This representation is useful since two marginal survival functions  $S_1$  and  $S_2$  are separated from the dependence structure  $C$ . Under the survival copula model (2), the copula  $C$  is a bivariate distribution function for  $U = S_1(T_1)$  and  $V = S_2(T_2)$ , namely  $C(u, v) = \mathbb{P}(U \leq u, V \leq v)$ . Note that the survival copula model (2) gives a different model from the copula model  $\mathbb{P}(T_1 \leq t_1, T_2 \leq t_2) = C(1 - S_1(t_1), 1 - S_2(t_2))$  unless the copula is radially symmetric, i.e.,  $C(u, v) = u + v - 1 + C(1 - u, 1 - v)$ .

We will consider the following well-known families of bivariate copulas.

1. The independence copula:

$$C(u, v) = uv.$$

2. The Clayton copula [28]:

$$C_\theta(u, v) = \max\left((u^{-\theta} + v^{-\theta} - 1)^{-1/\theta}, 0\right), \quad \theta \in [-1, \infty) \setminus \{0\}.$$

3. The Gumbel copula [29]:

$$C_\theta(u, v) = \exp\left\{-\left[(-\log u)^{\theta+1} + (-\log v)^{\theta+1}\right]^{\frac{1}{\theta+1}}\right\}, \quad \theta \in [0, \infty).$$

4. The Frank copula [30]:

$$C_\theta(u, v) = -\frac{1}{\theta} \log\left(1 + \frac{(e^{-\theta u} - 1)(e^{-\theta v} - 1)}{e^{-\theta} - 1}\right), \quad \theta \in (-\infty, \infty) \setminus \{0\}.$$

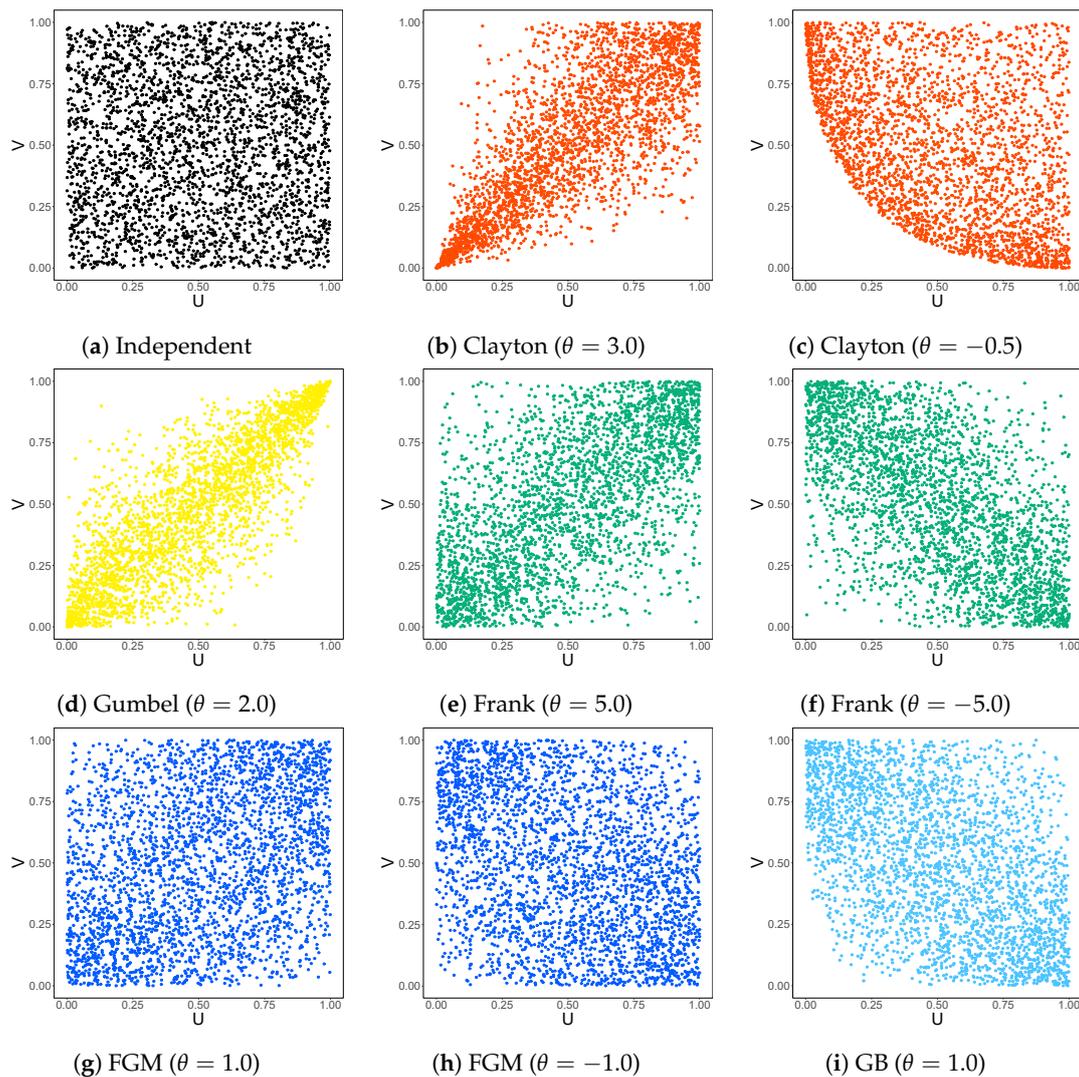
5. The Farlie–Gumbel–Morgenstern (FGM) copula [31]:

$$C_\theta(u, v) = uv + \theta uv(1 - u)(1 - v), \quad \theta \in [-1, 1].$$

6. The Gumbel–Barnett (GB) copula [24,32,33]:

$$C_{\theta}(u, v) = uv \exp(-\theta \log u \log v), \quad \theta \in [0, 1].$$

In Figure 1, we present scatter plots for  $(U, V)$  generated from various copulas with selected values of  $\theta$ . The Clayton copula (Figure 1b,c) shows lower tail dependence, the Gumbel copula (Figure 1d) shows upper tail dependence, Frank copula (Figure 1e,f) shows symmetric dependence around the median, and the FGM copula (Figure 1g,h) is similar to the Frank copula; both copulas are radially symmetric. Unlike other copulas, the GB copula exhibits negative dependence only (Figure 1i).



**Figure 1.** Scatter plots of 3000 data points generated from the copula distribution with parameter  $\theta$ .

These copulas have been applied to survival data and other data analyses. The Clayton copula was applied to survival data with dependent censoring. For instance, Schneider et al. [34] modeled dependence between survival and dependent censoring times in the survival data of tuberculosis cure. The Clayton, Gumbel, and Frank copulas were also applied to dependently censored data in clinical trials or observational studies [35–37]. The Gumbel, Frank, and FGM copulas were often used in competing risks models on survival data analysis [26,38–40]. Copulas were also applied to multivariate meta-analysis; Shih et al. and Shih et al. [41,42] proposed bivariate Clayton, FGM, and Gumbel models for bivariate meta-analysis. The Gumbel–Barnett copula has the simple form and is suitable

for modeling negative dependence [24,32]. Therefore, it is important to consider a variety of copulas for dependent survival times.

To see the strength of dependence in a copula,  $\theta$  can be transformed to Kendall’s  $\tau$ . Kendall’s  $\tau$  is a well-known measure to assess the dependence between two variables [20,24]. Kendall’s  $\tau$  does not depend on the marginals, and is solely determined by the copula (by  $\theta$ ). Appendix A provides the formulas of Kendall’s  $\tau$  for the Clayton, Gumbel, Frank, FGM, and GB copulas. Therefore, Kendall’s  $\tau$  is advantageous over the Pearson correlation for  $T_1$  and  $T_2$ .

### 2.2. Proposed Method for Computing $p$

In this section, we propose a new formula for computing  $p$  under the survival copula model (2). Let  $U = S_1(T_1)$  and  $V = S_2(T_2)$ . For computing  $p$ , we will use the conditional distribution function for  $U$  given  $V = v$ , which is the partial derivative of  $C$  with respect to  $v$ :

$$C^{[0,1]}(u, v) = \mathbb{P}(U \leq u \mid V = v) = \frac{\partial C(u, v)}{\partial v}.$$

Then, by slightly modifying the theorem given by Emura & Pan [22],  $p$  can be expressed as the univariate integral on  $[0, 1]$ :

$$\begin{aligned} p &= \mathbb{P}(T_1 > T_2) + \frac{1}{2}\mathbb{P}(T_1 = T_2) \\ &= \mathbb{P}(S_1(T_1) < S_1(T_2)) = \mathbb{P}(U < S_1(S_2^{-1}(V))) = \mathbb{E}[\mathbb{P}(U < S_1(S_2^{-1}(V)) \mid V)] \quad (3) \\ &= \int_0^1 C^{[0,1]}(S_1(S_2^{-1}(v)), v)dv. \end{aligned}$$

We note that the theorem of Emura & Pan [22] is not directly applicable to the survival copula model (2) since that theorem is designed for the copula model  $\mathbb{P}(T_1 \leq t_1, T_2 \leq t_2) = C(1 - S_1(t_1), 1 - S_2(t_2))$ .

In order to compute  $p$  by the above formulas, we need to specify  $S_1$ ,  $S_2$ , and  $\theta$ . One can specify  $S_1$  and  $S_2$  by continuous parametric models that will be discussed in Section 2.4. One can try different values for  $\theta$  in a sensitivity analysis. Note that the above calculations are not applicable for discrete parametric models for  $S_1$  and  $S_2$ .

### 2.3. Computing $p$ with Follow-Up Time

In this section, we assume that every subject has a common follow-up time  $\tau > 0$ . For survival data, the follow-up period is often limited. When a subject survives longer than the follow-up period, one may treat the survival time of the subject as equal to the follow-up period [6,43,44]. This means that we define the Mann–Whitney effect for  $\min(T_1, \tau)$  and  $\min(T_2, \tau)$ . We now obtain  $p$  with the follow-up time  $\tau$  from the following theorem, a straightforward expansion of the theorem of Emura & Pan [22].

**Theorem 1.** *The Mann–Whitney effect  $p$  with a follow-up time  $\tau$  is written as the univariate integral:*

$$\begin{aligned} p_\tau &= \mathbb{P}(\min(T_1, \tau) > \min(T_2, \tau)) + \frac{1}{2}\mathbb{P}(\min(T_1, \tau) = \min(T_2, \tau)) \\ &= \mathbb{P}(T_1 > T_2, T_2 < \tau) + \frac{1}{2}\mathbb{P}(T_1 > \tau, T_2 > \tau) \\ &= \mathbb{P}(U < S_1(S_2^{-1}(V)), V > S_2(\tau)) + \frac{1}{2}C(S_1(\tau), S_2(\tau)) \quad (4) \\ &= \int_{S_2(\tau)}^1 C^{[0,1]}(S_1(S_2^{-1}(v)), v)dv + \frac{1}{2}C(S_1(\tau), S_2(\tau)). \end{aligned}$$

Furthermore,  $p_\tau$  tends to  $p$  as  $\tau \rightarrow \infty$ . That is,

$$\lim_{\tau \rightarrow \infty} p_\tau = \int_0^1 C^{[0,1]}(S_1(S_2^{-1}(v)), v)dv + \frac{1}{2}C(0,0) = p.$$

Appendix B provides the formulas of  $p$  and  $p_\tau$  for the Clayton, Gumbel, Frank, FGM, and GB copulas.

### 2.4. Marginal Survival Distributions

To compute  $p$  and  $p_\tau$ , we considered the following parametric marginal distributions for the group  $j \in \{1, 2\}$ :

1. The exponential distribution:

$$S_j(t) = \exp(-\lambda_j t), \quad \lambda_j > 0,$$

$$S_i(S_j^{-1}(v)) = v^{\frac{\lambda_i}{\lambda_j}},$$

where  $\lambda_j$  is a rate parameter.

2. The Weibull distribution:

$$S_j(t) = \exp(-\lambda_j t^{k_j}), \quad \lambda_j > 0, k_j > 0,$$

$$S_i(S_j^{-1}(v)) = \exp\left(-\lambda_i \left(-\frac{\log v}{\lambda_j}\right)^{\frac{k_i}{k_j}}\right),$$

where  $\lambda_j$  is a scale parameter and  $k_j$  is a shape parameter.

3. The gamma distribution:

$$S_j(t) = 1 - \frac{\gamma(k_j, \lambda_j t)}{\Gamma(k_j)}, \quad \lambda_j > 0, k_j > 0,$$

where  $\lambda_j$  is a scale parameter and  $k_j$  is a shape parameter, and  $\Gamma(k)$  is the gamma function, and  $\gamma(k, \lambda t) = \int_0^{\lambda t} x^{k-1} e^{-x} dx$  is the lower incomplete gamma function. The gamma distribution has no simple closed-form expression for the inverse survival function. Therefore, one can use approximations for the inverse survival function. In this article, we use the R software version 4.2.3 function “*qgamma*” to calculate  $S_j^{-1}(v)$ .

4. The log-normal distribution:

$$S_j(t) = \frac{1}{\sqrt{2\pi\sigma_j^2}} \int_t^\infty \frac{1}{y} \exp\left\{-\frac{1}{2\sigma_j^2}(\log y - \mu_j)^2\right\} dy,$$

where  $\mu_j$  is a location parameter and  $\sigma_j$  is a scale parameter. Note that  $S_j(t)$  has no simple closed-form expression for the inverse survival function. Therefore, one can use approximations for the inverse survival function. In this article, we use the R software function “*qlnorm*” to calculate  $S_j^{-1}(v)$ .

5. The Burr III distribution:

$$S_j(t) = 1 - (1 + t^{-c_j})^{-k_j}, \quad c_j > 0, k_j > 0,$$

$$S_i(S_j^{-1}(v)) = 1 - \left(1 + \left((1 - v)^{-\frac{1}{k_j}} - 1\right)^{\frac{c_i}{c_j}}\right)^{-k_i},$$

where  $c_j, k_j$  are shape parameters.

In Figure 2, we present survival curves of the exponential, Weibull, gamma, log-normal, and Burr III distributions with different parameters. These plots show that these distributions can represent almost any continuous survival curve that will be encountered in practice.

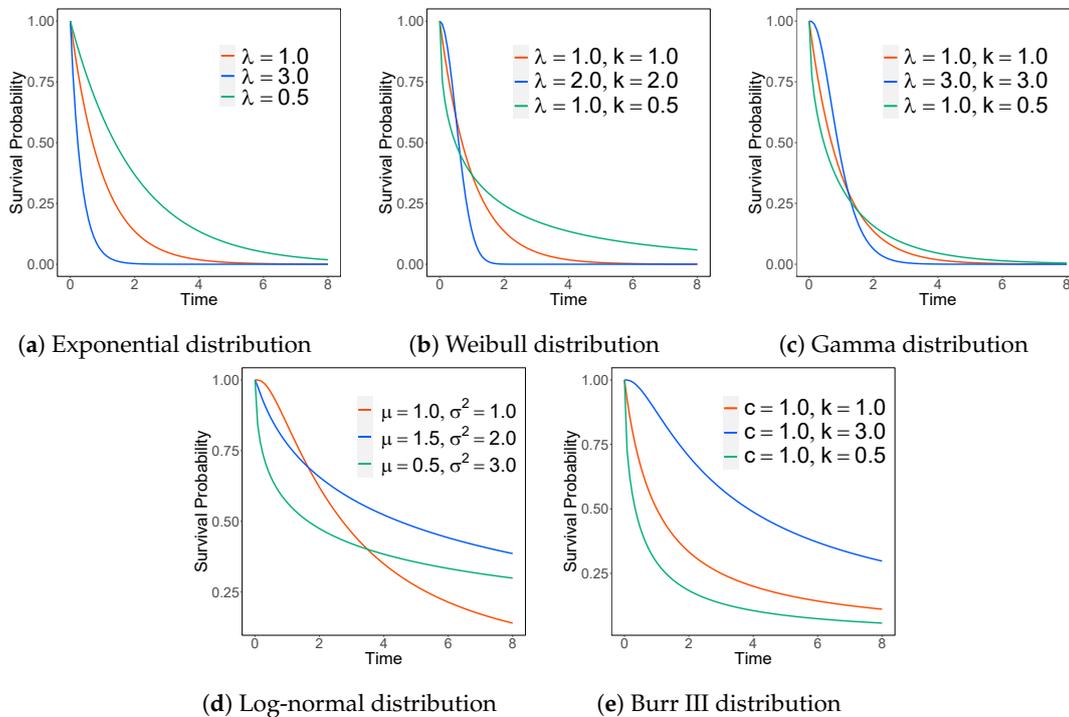


Figure 2. The plots for parametric survival functions.

**Example 1.** Let the marginals  $S_1(t), S_2(t)$  be the exponential distributions with parameter  $\lambda_1 = 1, \lambda_2 = 2$ . Assume that  $T_1$  and  $T_2$  are independent. Then, by Theorem 1 with  $\tau = \infty$ ,  $p$  is given by

$$\begin{aligned}
 p &= \int_0^1 C_{\theta}^{[0,1]}(S_1(S_2^{-1}(v)), v) dv \\
 &= \int_0^1 \exp\left(\frac{\lambda_1}{\lambda_2} \log v\right) dv = \frac{\lambda_2}{\lambda_1 + \lambda_2} = \frac{2}{3}.
 \end{aligned}$$

**Example 2.** Let the marginals  $S_1(t), S_2(t)$  be the exponential distributions with parameter  $\lambda_1 = 1, \lambda_2 = 2$ . Assume that  $C_{\theta}(u, v)$  is the Clayton copula with parameter  $\theta = 3$ . Then, Kendall's  $\tau$  is given by

$$\text{Kendall's } \tau = \frac{\theta}{\theta + 2} = 0.6$$

and by Theorem 1 with  $\tau = \infty$ ,  $p$  is given by

$$\begin{aligned}
 p &= \int_0^1 C_{\theta}^{[0,1]}(S_1(S_2^{-1}(v)), v) dv \\
 &= \int_0^1 v^{-\theta-1} \left(v^{-\theta \frac{\lambda_1}{\lambda_2}} + v^{-\theta} - 1\right)^{-\frac{1}{\theta}-1} dv \\
 &= \int_0^1 v^{-4} \left(v^{-\frac{3}{2}} + v^{-3} - 1\right)^{-\frac{4}{3}} dv = 0.84.
 \end{aligned}$$

When  $\tau < \infty$ ,  $p_\tau$  is given by

$$p_\tau = \int_0^1 C^{[0,1]}(S_1(S_2^{-1}(v)), v) dv + \frac{1}{2} \mathbb{P}(\min(T_1, \tau) = \min(T_2, \tau))$$

$$= \int_{e^{2\tau}}^1 v^{-4} \left( v^{-\frac{3}{2}} + v^{-3} - 1 \right)^{-\frac{4}{3}} dv + \frac{1}{2} \left( e^{3\tau} + e^{6\tau} - 1 \right)^{-\frac{1}{3}}.$$

When  $\tau = 0.5$ ,  $p_\tau = 0.68$ . Here, we computed the last two equations by a numerical integration by the R software version 4.2.3 function “integrate”.

Parameters for the marginal distributions can be estimated by maximum likelihood estimators (MLEs) when the survival data are available in two groups (Section 5).

### 2.5. Sensitivity Analysis by Copulas

To comprehensively cover the possible dependence structures between  $T_1$  and  $T_2$ , we select the Clayton, Gumbel, Frank, FGM, and GB copulas, and set their parameters. Table 1 contains positive dependence and negative dependence with weak and strong correlations, as well as the independence. However, the selection of an appropriate copula may be difficult unless both  $T_1$  and  $T_2$  are observed for the same subject (e.g., by cross-over designs). This is the case for the data examples (Section 5), where  $T_1$  and  $T_2$  are observed for different subjects. Let  $Z$  be the group indicator ( $Z = 1$  when  $T_1$  is observed;  $Z = 2$  when  $T_2$  is observed). What we observe is one of  $T_1$  and  $T_2$ , namely,  $T = T_1 \mathbb{1}(Z = 1) + T_2 \mathbb{1}(Z = 2)$ . As  $T_1$  and  $T_2$  are never observed simultaneously, the copula for  $T_1$  and  $T_2$  are not identifiable [18,19]. This means that copula model selection and goodness-of-fit test are not feasible. Therefore, as one cannot specify a single copula, we suggest computing  $p$  under all these copulas and dependence parameters to see how  $p$  changes under various dependence. Such sensitivity analyses allow researchers to obtain the bounds of  $p$  under the possible dependence structures.

**Table 1.** Selection of copulas and their parameters.

Copula	$\theta$	Kendall's $\tau$
Clayton	1.0	0.33
	5.0	0.71
	10.0	0.83
Gumbel	0.0	0.00
	4.0	0.80
Frank	−20.0	−0.82
	−5.0	−0.46
	1.0	0.11
	5.0	0.46
FGM	−1.0	−0.22
	0.0	0.00
GB	1.0	0.22
	0.5	−0.21
	1.0	−0.36

### 3. Software and Web App

We developed a Shiny-based web app to implement the proposed method for computing  $p$  and  $p_\tau$ . The app is available at ([https://nkosuke.shinyapps.io/shiny\\_survival/](https://nkosuke.shinyapps.io/shiny_survival/) accessed on 7 April 2024) and can be used in any environment, including smartphones. Using this app, users can choose a marginal survival distribution, copula, and the relevant parameters to compute  $p$  and  $p_\tau$ . Our app is easy to use without knowledge of the R software version 4.2.3. The app works without data, because it does not estimate the parameters of the marginals and copula. All relevant parameters are entered by users. The

app can compute  $p$  and  $p_\tau$  across different parametric distributions (Section 2.4) and copula models (Section 2.5).

### 3.1. Input

The web app requires users to select input values to compute  $p$  and  $p_\tau$ . The left panel of Figure 3 shows input icons, where users can select marginal distributions from the exponential, Weibull, gamma, log-normal, and Burr III distributions and a copula from the Clayton, Gumbel, Frank, FGM, and GB copulas. One can choose marginal distributions and copulas, set their parameters, and choose the language displayed on the screen in the input panels on the left-hand of the app (Figure 3). Furthermore, users can set a follow-up time  $\tau$ .

## Select and Display Results

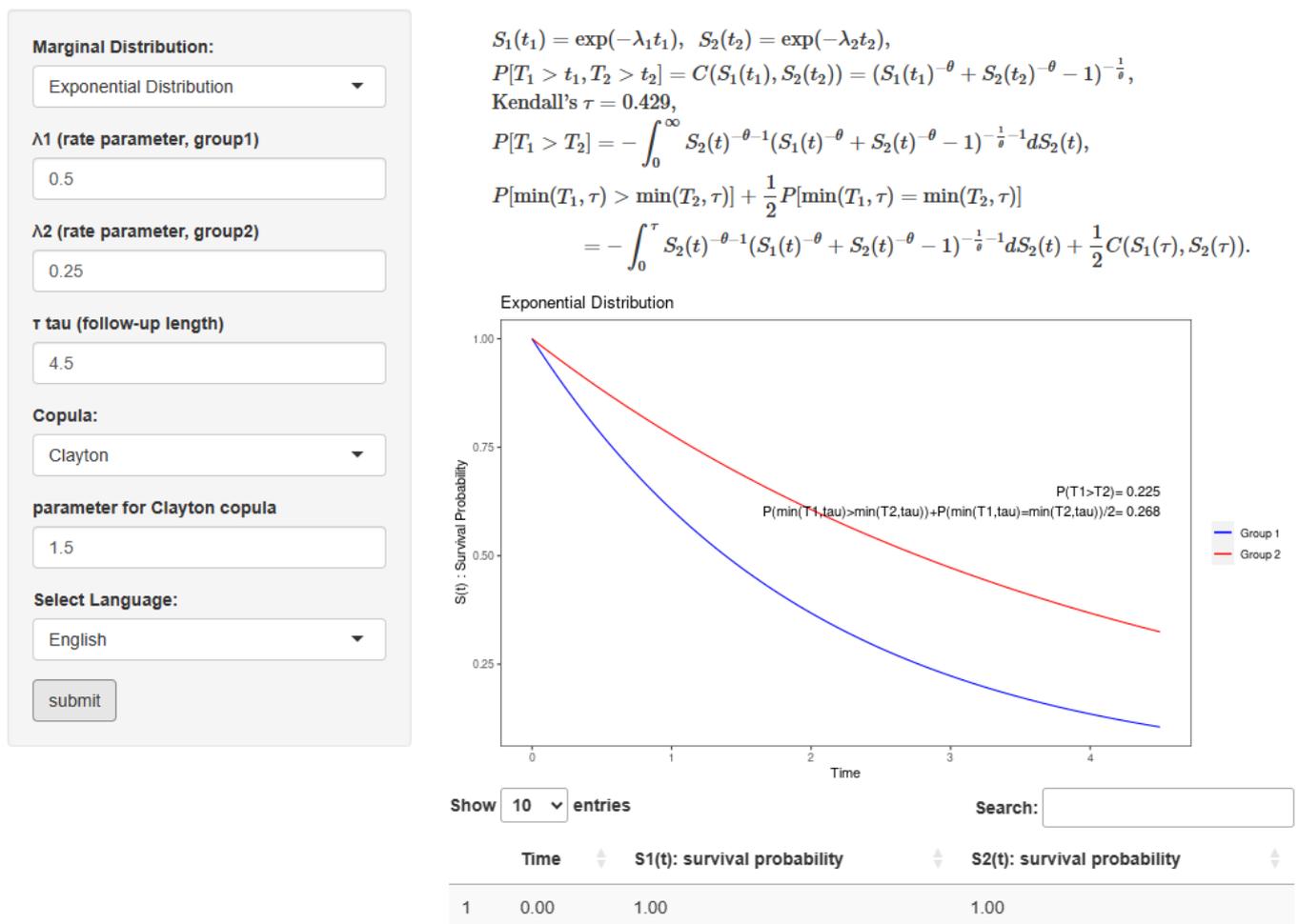


Figure 3. The web app showing the results for computing  $p$  and  $p_\tau$  under the exponential distribution.

### 3.2. Output

This app displays key formulas, survival curves of two groups, the values of  $p$  and  $p_\tau$ , and the value of Kendall's  $\tau$ . These formulas include marginal and bivariate survival functions and formulas of  $p$  and  $p_\tau$  based on the input values. The theoretical values of  $p$  and  $p_\tau$  are displayed together with survival curves.

### 3.3. Example of Using the App

Users may set the marginal survival distribution = "Exponential Distribution",  $\lambda_1 = 0.5$ ,  $\lambda_2 = 0.25$ ,  $\tau = 4.5$ , copula = "Clayton",  $\theta = 1.5$  and the language displayed on the screen "English", and push "submit" button, and then they can obtain  $p = 0.225$ ,  $p_\tau = 0.268$ .



Table 2. Cont.

Distribution	Copula	$\theta$	Kendall's $\tau$	$\tau = 0.5$		$\tau = 2$		$\tau = 5$		$\tau = \infty$		
				$P_{\tau, \text{theory}}$	$P_{\tau, \text{sim}}$							
Gamma	Clayton	1.0	0.33	0.529	0.529	0.651	0.649	0.679	0.678	0.679	0.680	
		5.0	0.71	0.530	0.528	0.763	0.763	0.809	0.810	0.809	0.810	
		10.0	0.83	0.534	0.533	0.817	0.816	0.862	0.863	0.863	0.863	
	Gumbel	0.0	0.00	0.530	0.530	0.611	0.612	0.615	0.614	0.615	0.616	
		4.0	0.80	0.545	0.546	0.813	0.813	0.853	0.853	0.854	0.853	
	Frank	-5.0	-0.46	0.532	0.530	0.584	0.584	0.584	0.583	0.584	0.584	
		1.0	0.11	0.530	0.529	0.622	0.622	0.628	0.628	0.628	0.628	
		5.0	0.46	0.530	0.530	0.679	0.679	0.694	0.692	0.694	0.697	
	FGM	-1.0	-0.22	0.531	0.533	0.591	0.591	0.592	0.591	0.592	0.592	
		0.0	0.00	0.530	0.529	0.611	0.610	0.615	0.614	0.615	0.617	
		1.0	0.22	0.529	0.529	0.631	0.631	0.639	0.640	0.639	0.641	
	GB	0.5	-0.21	0.531	0.532	0.597	0.598	0.598	0.598	0.598	0.597	
		1.0	-0.36	0.532	0.532	0.589	0.592	0.590	0.588	0.589	0.588	
	Log-normal	Clayton	1.0	0.33	0.580	0.582	0.596	0.594	0.593	0.591	0.573	0.574
			5.0	0.71	0.599	0.599	0.658	0.658	0.675	0.677	0.619	0.619
			10.0	0.83	0.618	0.617	0.715	0.715	0.756	0.755	0.684	0.684
		Gumbel	0.0	0.00	0.574	0.575	0.576	0.578	0.569	0.570	0.564	0.562
			4.0	0.80	0.652	0.652	0.745	0.746	0.760	0.759	0.728	0.726
Frank		-5.0	-0.46	0.567	0.567	0.553	0.554	0.547	0.547	0.547	0.544	
		1.0	0.11	0.577	0.576	0.584	0.586	0.578	0.580	0.571	0.570	
		5.0	0.46	0.593	0.593	0.624	0.625	0.623	0.621	0.609	0.607	
FGM		-1.0	-0.22	0.568	0.567	0.560	0.561	0.553	0.553	0.550	0.551	
		0.0	0.00	0.574	0.574	0.576	0.576	0.569	0.569	0.564	0.560	
		1.0	0.22	0.580	0.580	0.591	0.593	0.585	0.586	0.577	0.579	
GB		0.5	-0.21	0.571	0.572	0.563	0.566	0.558	0.555	0.558	0.557	
		1.0	-0.36	0.568	0.567	0.557	0.557	0.550	0.554	0.549	0.550	
Burr III		Clayton	1.0	0.33	0.661	0.662	0.753	0.753	0.760	0.761	0.747	0.748
			5.0	0.71	0.665	0.665	0.819	0.820	0.883	0.883	0.863	0.865
			10.0	0.83	0.666	0.666	0.832	0.832	0.913	0.914	0.896	0.897
		Gumbel	0.0	0.00	0.660	0.661	0.714	0.715	0.701	0.701	0.697	0.696
			4.0	0.80	0.667	0.665	0.832	0.832	0.885	0.883	0.871	0.871
	Frank	-5.0	-0.46	0.658	0.658	0.663	0.662	0.650	0.649	0.650	0.653	
		1.0	0.11	0.661	0.660	0.730	0.731	0.720	0.719	0.715	0.716	
		5.0	0.46	0.664	0.663	0.788	0.789	0.798	0.797	0.789	0.788	
	FGM	-1.0	-0.22	0.658	0.658	0.683	0.683	0.666	0.666	0.665	0.664	
		0.0	0.00	0.660	0.660	0.714	0.711	0.701	0.701	0.697	0.699	
		1.0	0.22	0.661	0.662	0.744	0.747	0.737	0.736	0.730	0.729	
	GB	0.5	-0.21	0.659	0.659	0.692	0.693	0.676	0.677	0.675	0.676	
		1.0	-0.36	0.658	0.658	0.673	0.671	0.659	0.659	0.659	0.659	

Note: We set the exponential distributions with  $(\lambda_1, \lambda_2) = (1.0, 2.0)$ , the Weibull distributions with  $(\lambda_1, k_1, \lambda_2, k_2) = (1.0, 0.5, 2.0, 1.0)$ , the gamma distributions with  $(\lambda_1, k_1, \lambda_2, k_2) = (1.0, 1.5, 2.0, 2.0)$ , the log-normal distributions with  $(\mu_1, \sigma_1^2, \mu_2, \sigma_2^2) = (0.7, 1.5, 0.3, 2.0)$ , and the Burr III distributions with  $(c_1, k_1, c_2, k_2) = (1.5, 3.0, 1.0, 1.0)$ .

### 5. Data Analysis

In this section, we apply our proposed methods to a tongue cancer dataset and a prostate cancer dataset. The two datasets are publicly available and anonymized. The purpose of the data analyses is to show how the proposed methods are implemented and how the results offer new insights beyond Efron's traditional estimator for the Mann-Whitney effect.

Before analyzing the real datasets, we introduce basic notations and ideas for estimating  $p$  by using censored data. We consider survival times  $(T_{i1}, T_{i2})$  for two groups, a group indicator  $Z_i$  ( $Z_i = 1$  for group 1 and  $Z_i = 2$  for group 2), and censoring time  $C_i$  for  $i = 1, \dots, n$ . The sample size for each group is  $n_j = \sum_i \mathbb{1}(Z_i = j)$ ,  $j = 1, 2$ . With  $T_i = T_{i1} \mathbb{1}(Z_i = 1) + T_{i2} \mathbb{1}(Z_i = 2)$ , we observe is  $X_i = \min(T_i, C_i)$ ,  $\Delta_i = \mathbb{1}(T_i \leq C_i)$ , and  $Z_i$  for  $i = 1, \dots, n$ . In this observation, only one of  $T_{i1}$ ,  $T_{i2}$ , and  $C_i$  is observed ( $T_{i1}$  and  $T_{i2}$  are never observed simultaneously) for each  $i$ , making it difficult to identify a copula for  $(T_{i1}, T_{i2})$ . Without specifying a copula for possibly dependent survival times  $(T_{i1}, T_{i2})$ , one can still estimate the marginal distributions. Using the exponential distribution, we obtained the MLE of the exponential hazard rate  $\hat{\lambda}_j$ ,  $j = 1, 2$  by

$$\hat{\lambda}_j = \frac{\sum_{i=1}^n \Delta_i \mathbb{1}(Z_i = j)}{\sum_{i=1}^n X_i \mathbb{1}(Z_i = j)}, j = 1, 2.$$

Then, by applying the values of the MLE to the proposed Shiny web app, we obtained the estimators  $\hat{p}$  and  $\hat{p}_\tau$ , where  $\tau$  was chosen appropriately (Sections 5.1 and 5.2). The estimators were computed under various copulas as one cannot specify a single copula (see

Section 2.5). On the other hand, under the independence assumption of  $T_{i1}$  and  $T_{i2}$ , Efron’s estimator [1] of  $p$  is

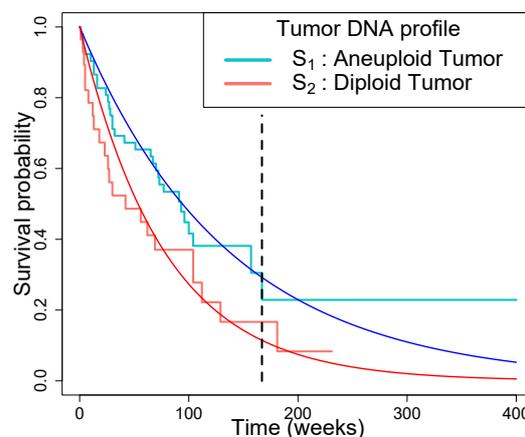
$$\hat{p}^{KM} = - \int_0^\infty \hat{S}_1^\pm(t) d\hat{S}_2(t),$$

$$\hat{p}_\tau^{KM} = - \int_0^{\hat{S}_2^{-1}(\tau)} \hat{S}_1^\pm(t) d\hat{S}_2(t) + \frac{1}{2} \hat{S}_1^\pm(\tau) \hat{S}_2^\pm(\tau),$$

where  $\hat{S}_j^\pm(t) = [\hat{S}_j(t+) + \hat{S}_j(t-)]/2$  and  $\hat{S}_j(t)$  is a Kaplan–Meier (KM) estimator for group  $j = 1, 2$ . However, these benchmark estimates are subject to the independence of two groups. Therefore, the proposed estimator is useful to examine the sensitivity under a variety of dependence structures via copulas. We apply a jackknife estimator of standard error (SE) to measure the estimators’ uncertainty, and employ the normal approximation to obtain  $p$ -values for testing  $H_0 : p_\tau = 1/2$  vs.  $H_1 : p_\tau \neq 1/2$ .

### 5.1. Tongue Cancer Data

The tongue dataset is available in the R software version 4.2.3 package *KMsurv*. It has 80 observations and contains: type (tumor DNA profile: 1 = aneuploid tumor, 2 = diploid tumor), time (time to death or on-study time (weeks)), and death (event indicator: 0 = alive, 1 = dead). It contains  $n_1 = 52$  observations in the DNA-aneuploid tumor group ( $j = 1$ ), and  $n_2 = 28$  observations in the DNA-diploid tumor group ( $j = 2$ ). We considered the follow-up time  $\tau = \min\{\max_i X_i \Delta_i \mathbb{1}(Z_i = 1), \max_i X_i \Delta_i \mathbb{1}(Z_i = 2)\}$  and obtained  $\tau = 167$ . The tongue cancer data resulted in  $\hat{\lambda}_1 = 0.00736$ ,  $\hat{\lambda}_2 = 0.0130$  and  $\hat{p}_\tau^{KM} = 0.615$ . In Figure 4, the KM estimators of each group and the estimated exponential survival curves are plotted.



**Figure 4.** KM estimators for the DNA-aneuploid tumor and the DNA-diploid tumor group and the exponential survival curves with the MLEs of exponential hazard rate (darker blue and red lines),  $\hat{\lambda}_1 = 0.00736$ ,  $\hat{\lambda}_2 = 0.0130$ . The vertical line signifies the follow-up time  $\tau = 167$ .

We conducted sensitivity analyses using a copula-based approach (see Section 2.5). We calculated  $\hat{p}$  and  $\hat{p}_\tau$  by Theorem 1 under weak, strong positive, and negative independence. We calculated  $\hat{p}_\tau$  via the web app (Figure 5). Table 3 shows the output under the independent, Clayton, Gumbel, Frank, FGM, and GB copula with parameter  $\theta \in \{1, 5\}$  (the Clayton),  $\theta = 4$  (the Gumbel),  $\theta \in \{-20, -5, 5\}$  (the Frank),  $\theta \in \{-1, 1\}$  (the FGM),  $\theta \in \{0.5, 1\}$  (the GB). The results for all scenarios are summarized in Table 3. We obtained the  $\hat{p}_\tau$  ranging from 0.596 to 0.895, which is equivalent to Cohen’s  $d$  being greater than or equal to 0.5. Therefore, we concluded that subjects in the DNA-aneuploid tumor group survive longer than in those the DNA-diploid tumor group. This conclusion did not change under any of the conducted dependence structures.

### Select and Display Results

**Marginal Distribution:**  
 Exponential Distribution

**λ1 (rate parameter, group1)**  
 0.00736

**λ2 (rate parameter, group2)**  
 0.0130

**τ tau (follow-up length)**  
 167

**Copula:**  
 Clayton

**parameter for Clayton copula**  
 1.0

**Select Language:**  
 English

submit

$$S_1(t_1) = \exp(-\lambda_1 t_1), \quad S_2(t_2) = \exp(-\lambda_2 t_2),$$

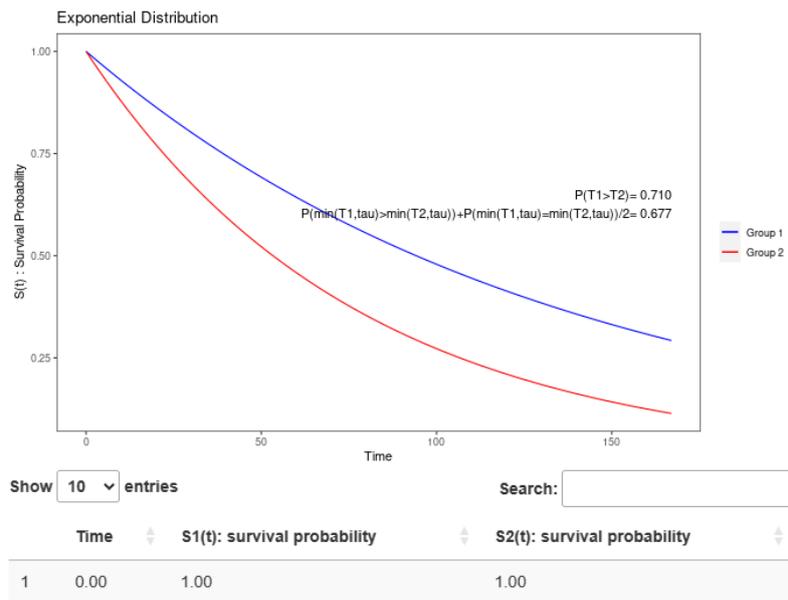
$$P[T_1 > t_1, T_2 > t_2] = C(S_1(t_1), S_2(t_2)) = (S_1(t_1)^{-\theta} + S_2(t_2)^{-\theta} - 1)^{-\frac{1}{\theta}},$$

Kendall's  $\tau = 0.333$ ,

$$P[T_1 > T_2] = - \int_0^{\infty} S_2(t)^{-\theta-1} (S_1(t)^{-\theta} + S_2(t)^{-\theta} - 1)^{-\frac{1}{\theta}-1} dS_2(t),$$

$$P[\min(T_1, \tau) > \min(T_2, \tau)] + \frac{1}{2} P[\min(T_1, \tau) = \min(T_2, \tau)]$$

$$= - \int_0^{\tau} S_2(t)^{-\theta-1} (S_1(t)^{-\theta} + S_2(t)^{-\theta} - 1)^{-\frac{1}{\theta}-1} dS_2(t) + \frac{1}{2} C(S_1(\tau), S_2(\tau)).$$



**Figure 5.** Example for the tongue cancer dataset on the app. This setting is marginal distribution: “Exponential”,  $\lambda_1 = 0.00736, \lambda_2 = 0.0130, \tau = 167$ , copula: “Clayton”, copula parameter:  $\theta = 1$ , and language: “English”.

**Table 3.** Estimates  $\hat{p}_\tau (\tau = 167)$  for fitting the KM estimator (independent) and with exponential marginal survival distributions (the independent, Clayton, Gumbel, Frank, FGM, GB copulas) for the tongue cancer dataset.

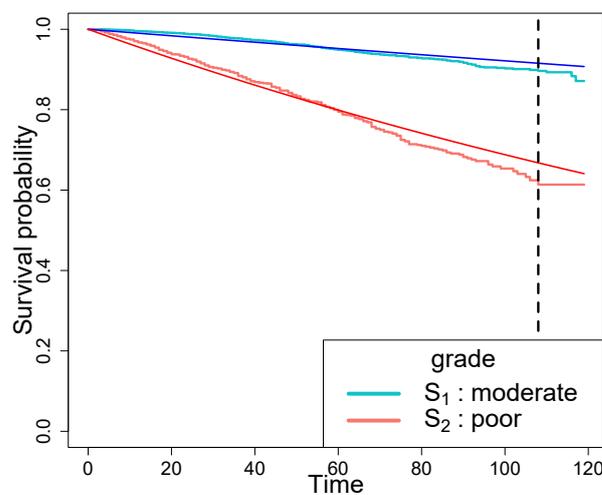
Copula	Marginal Distribution	$\theta$	$\hat{p}$	SE	p-Value	$\hat{p}_\tau (\tau = 167)$	SE	p-Value
Independent	KM estimator	-	-	-	-	0.624	0.071	0.079
Independent	exponential	-	0.638	0.076	0.070	0.633	0.075	0.076
Clayton	exponential	1.0	0.709	0.096	0.029	0.676	0.096	0.067
		5.0	0.856	0.075	<0.001	0.799	0.100	0.003
Gumbel	exponential	4.0	0.944	0.084	<0.001	0.895	0.095	<0.001
Frank	exponential	-20.0	0.596	0.055	0.080	0.596	0.055	0.080
		-5.0	0.600	0.057	0.081	0.600	0.057	0.081
		5.0	0.733	0.111	0.036	0.714	0.108	0.046
FGM	exponential	-1.0	0.609	0.063	0.082	0.609	0.063	0.083
		1.0	0.666	0.090	0.063	0.658	0.088	0.072
GB	exponential	0.5	0.617	0.066	0.077	0.617	0.066	0.078
		1.0	0.606	0.060	0.079	0.606	0.060	0.080

### 5.2. Prostate Cancer Data

The prostate cancer data are available in the R software version 4.2.3 package *asaur* [45]. They have 14,294 observations and contain grade (moderately differentiated and poorly differentiated), survTime (time from diagnosis to death or last date known alive), and status (event indicator: 0 = censored, 1 = death from prostate cancer). They contain  $n_1 = 10,988$

observations in the moderately differentiated group ( $j = 1$ ), and  $n_2 = 3306$  observations in the poorly differentiated group ( $j = 2$ ).

The prostate cancer data resulted in  $\hat{\lambda}_1 = 0.000817, \hat{\lambda}_2 = 0.00374, \tau = 108$ , and  $\hat{p}^{KM} = 0.679$ . In Figure 6, we plot the KM estimators of each group and estimated exponential survival curves. We calculated  $\hat{p}_\tau$  by Theorem 1 with copulas and several parameters. We calculated  $\hat{p}_\tau$  under the independent, Clayton, Gumbel, Frank, FGM, and GB copulas with parameter  $\theta \in \{1, 5\}$  (the Clayton),  $\theta = 4$  (the Gumbel),  $\theta \in \{-20, -5, 5\}$  (the Frank),  $\theta \in \{-1, 1\}$  (the FGM),  $\theta \in \{0.5, 1\}$  (the GB) via the web app (Figure 7). The results for all scenarios are summarized in Table 4. We obtained the  $\hat{p}_\tau$  ranging from 0.623 to 0.666, which is equivalent to Cohen’s  $d$  being equal to 0.5. Therefore, we concluded that subjects in the moderately differentiated group survive longer than those in the poorly differentiated group. The range of  $\hat{p}_\tau$  is narrower than one of tongue cancer dataset, because the large difference of survival functions may not be influenced by copulas.

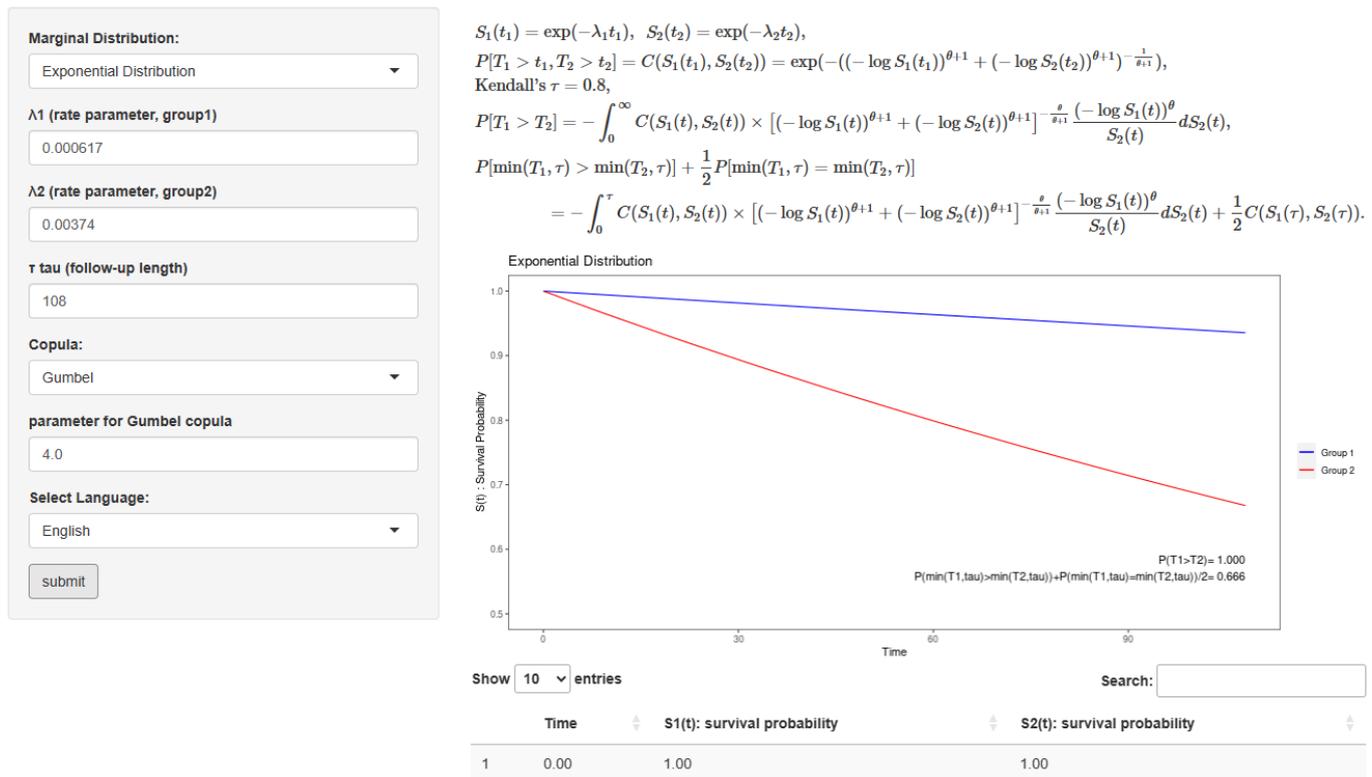


**Figure 6.** KM estimators for the moderately differentiated and the poorly differentiated group and exponential survival curves with the MLEs of exponential hazard rate (darker blue and red lines),  $\hat{\lambda}_1 = 0.000817, \hat{\lambda}_2 = 0.00374$ . The vertical line signifies the follow-up time  $\tau = 108$ .

**Table 4.** Estimates  $p_\tau(\tau = 108)$  for fitting the KM estimator (independent) and  $p_\tau(\tau = 108)$  with the exponential marginal survival distributions (the independent, Clayton, Gumbel, Frank, FGM, GB copulas) for the prostate cancer dataset.

Copula	Marginal Distribution	$\theta$	$\hat{p}$	SE	$p$ -Value	$\hat{p}_\tau$ ( $\tau = 108$ )	SE	$p$ -Value
Independent	KM estimator	-	-	-	-	0.635	0.013	<0.001
Independent	exponential	-	0.821	0.010	<0.001	0.625	0.007	<0.001
Clayton	exponential	1.0	0.889	0.008	<0.001	0.626	0.007	<0.001
		5.0	0.958	0.003	<0.001	0.635	0.007	<0.001
Gumbel	exponential	4.0	0.999	<0.001	<0.001	0.666	0.006	<0.001
		-20.0	0.741	0.009	<0.001	0.624	0.007	<0.001
Frank	exponential	-5.0	0.753	0.010	<0.001	0.624	0.007	<0.001
		5.0	0.924	0.007	<0.001	0.632	0.007	<0.001
FGM	exponential	-1.0	0.777	0.011	<0.001	0.623	0.007	<0.001
		1.0	0.865	0.010	<0.001	0.626	0.007	<0.001
GB	exponential	0.5	0.786	0.010	<0.001	0.624	0.007	<0.001
		1.0	0.764	0.010	<0.001	0.623	0.007	<0.001

### Select and Display Results



**Figure 7.** Example for the tongue cancer dataset on the app. This setting is marginal distribution: “Exponential”,  $\lambda_1 = 0.000817, \lambda_2 = 0.00374, \tau = 108,$  copula: “Gumbel”, copula parameter:  $\theta = 4,$  and language: “English”.

### 6. Conclusions

The Mann–Whitney effect has been widely used for survival analysis, which can provide a meaningful measure of treatment effects on survival outcomes. However, the Mann–Whitney effect may not be interpreted as the true treatment effect under the dependence of the two survival times. In this article, we proposed a parametric copula-based approach for estimating the Mann–Whitney effect  $p$  under dependence structures for two survival times. We derived the formulas of  $p$  under a variety of copulas and marginal survival functions and modified it by  $p_{\tau}$  with  $\tau$  as the follow-up time. We also introduced a web-based calculator for  $p$  and  $p_{\tau}$  for users. Our simulation studies demonstrate the correctness of the proposed calculator under a variety of parametric marginal survival distributions and copulas. The results of the data analyses show that the proposed method provides possible changes in  $p$  and  $p_{\tau}$  under various dependence and enables the examination of the sensitivity.

In the examples of real datasets, we obtained  $p_{\tau}$  under the Clayton, Gumbel, Frank, FGM, and GB copulas with varying parameters. The value of  $p_{\tau}$  ranged from 0.596 to 0.895 in the tongue cancer dataset and from 0.623 to 0.666 in the prostate cancer dataset. We obtained narrow ranges whose lower bounds did not include the null value of 1/2. The results show that, under a variety of dependence structures, the interpretation of the Mann–Whitney effect does not change. This result is consistent with previous studies showing that Hand’s paradox does not occur under a strictly monotonic effect [15,46]. The considered dependence may not affect decision making in clinical research or practice. Although more complex dependence structures with various copulas might be considered, the conclusion may not change significantly.

On computing the Mann–Whitney effect under dependence, there are several previous works (Table 5). However, these works focused on specific combinations of marginals and copulas. The advantage of our article is that it can be applied to various marginals and copulas, not just to a specific one. Our app allows readers to verify the behavior of the Mann–Whitney effect under various dependences.

**Table 5.** Recent works on the Mann–Whitney effect with the copula models and marginal models.

	Copula	Marginal Distribution
Domma & Giordano [47]	FGM, Generalized FGM, Frank	Burr III, Dagum, Singh–Maddala
Gao et al. [48]	Mixed (Clayton, Gumbel, Frank)	Empirical
de Andrade et al. [49]	Clayton, Gumbel, Frank, Gauss, Plackett	Weibull, Gamma, Log-normal, Dagum
Rathie et al. [50]	Frank	Dagum, Log-Dagum
James et al. [51]	FGM	Rayleigh
Shang & Yan [52]	Clayton	Weibull, Kumaraswamy
Lima et al. [53]	Clayton, Frank, Gumbel–Hougaard	Generalized extreme value, Weibull, gamma

The main limitation of the present article is that we only discussed the “parametric” approach. Because the proposed method uses extrapolated parameters to compute the Mann–Whitney effect, the result strongly depends on the model assumptions of the marginals and copulas and the characteristics of the extrapolated parameter. However, in practice, researchers may use the “semiparametric” or “nonparametric” approaches [17]. In future work, we will examine the method of computing  $p$  without parametric assumptions and expand it to nonparametric or semiparametric models such as Cox regression. The anticipated challenge of nonparametric approach is that the empirical copula, which is a nonparametric approach to copula models, is often nonsmooth and not a genuine copula. Another extension is to include covariates or secondary outcomes, including time-varying effects in the model, which helps obtain narrow bounds for treatment effects [54–56]. Another limitation is that only one-parameter bivariate copulas and noninformative censoring were implemented. Multiparameter or multivariate copulas deserve attention [24,32,57]. There are also several recent works on copula-based approaches for dependent censoring [58–60]. If censoring is not independent of survival, the usual MLE and KM estimators are biased (hence, the estimators in Section 5 are all biased). Furthermore, the proposed method for two-sample comparison may be extended to multigroup comparisons using the factorial designs by adopting the relative treatment effects of Brunner & Puri and Dobler & Pauly [43,61].

**Supplementary Materials:** The following supporting information can be downloaded at: <https://www.mdpi.com/article/10.3390/math12101453/s1>, the R code for simulations and data analyses.

**Author Contributions:** Conceptualization, K.N. and T.E.; Methodology, K.N. and T.E.; Software, K.N. and Y.-C.L.; Resources, G.-Y.L.; Writing—original draft, K.N., R.U. and T.E.; Visualization, K.N.; Supervision, T.E.; Funding acquisition, R.U. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research was funded by [JSPS KAKENHI] grant number [22K11948], grant number [20H04147], and grant number [23K20374].

**Data Availability Statement:** As all data in Section 5 use publicly available data, written informed consent was not required.

**Acknowledgments:** We thank the editor and three reviewers for their time to comments on our paper. The comments from the reviewers greatly helped improve the presentation of the paper.

**Conflicts of Interest:** The authors declare no conflicts of interest.

**Appendix A. Kendall’s  $\tau$**

Under the copula model (2), Kendall’s  $\tau$  for  $T_1$  and  $T_2$  is expressed as

$$\text{Kendall's } \tau = 4 \int_0^1 \int_0^1 C_\theta(u, v)C_\theta(du, dv) - 1.$$

Kendall’s  $\tau$  of each copula is expressed as follows.

1. The independence copula:

$$\text{Kendall's } \tau = 0.$$

2. The Clayton copula:

$$\text{Kendall's } \tau = \frac{\theta}{\theta + 2}, \quad \theta \in [-1, \infty) \setminus \{0\}.$$

3. The Gumbel copula:

$$\text{Kendall's } \tau = \frac{\theta}{\theta + 1}, \quad \theta \in [0, \infty).$$

4. The Frank copula:

$$\text{Kendall's } \tau = 1 - \frac{4}{\theta} + \frac{4D_\theta}{\theta}, \quad \text{where } D_\theta = \frac{1}{\theta} \int_0^\theta \frac{x}{\exp(x) - 1} dx, \quad \theta \in (-\infty, \infty) \setminus \{0\}.$$

5. The FGM copula:

$$\text{Kendall's } \tau = \frac{2}{9}\theta, \quad \theta \in [-1, 1].$$

6. The GB copula:

$$\text{Kendall's } \tau = 1 - \frac{4}{\theta} \int_0^1 t(1 - \theta \log t) \log(1 - \theta \log t) dt, \quad \theta \in [0, 1].$$

**Appendix B. Examples of  $p$  and  $p_\tau$  with Different Copulas**

Equation (3) with different copulas is computed by the following formulas:

1. The Clayton copula:

$$p = \int_0^1 v^{-\theta-1} \left( \{S_1(S_2^{-1}(v))\}^{-\theta} + v^{-\theta} - 1 \right)^{-\frac{1}{\theta}-1} dv.$$

2. The Gumbel copula:

$$p = \int_0^1 \left\{ \exp \left\{ - [ (-\log S_1(S_2^{-1}(v)))^{\theta+1} + (-\log v)^{\theta+1} ]^{\frac{1}{\theta+1}} \right\} \right. \\ \left. \times \left[ \left( -\log(S_1(S_2^{-1}(v))) \right)^{\theta+1} + (-\log(v))^{\theta+1} \right]^{-\frac{\theta}{\theta+1}} \frac{(-\log(v))^\theta}{v} \right\} dv.$$

3. The Frank copula:

$$p = \int_0^1 \frac{e^{-\theta v} \left( e^{-\theta S_1(S_2^{-1}(v))} - 1 \right)}{\left( e^{-\theta} - 1 \right) + \left( e^{-\theta S_1(S_2^{-1}(v))} - 1 \right) \left( e^{-\theta v} - 1 \right)} dv.$$

4. The FGM copula:

$$p = \int_0^1 \left\{ S_1(S_2^{-1}(v)) + \theta S_1(S_2^{-1}(v)) \left( 1 - S_1(S_2^{-1}(v)) \right) (1 - 2v) \right\} dv.$$

5. The GB copula:

$$p = \int_0^1 \left\{ S_1(S_2^{-1}(v)) \left( 1 - \theta \log S_1(S_2^{-1}(v)) \right) v^{-\theta \log S_1(S_2^{-1}(v))} \right\} dv.$$

Equation (4) with different copulas is computed by the following formulas:

1. The Clayton copula:

$$p_\tau = \int_{S_2(\tau)}^1 \left\{ v^{-\theta-1} \left( \{S_1(S_2^{-1}(v))\}^{-\theta} + v^{-\theta} - 1 \right)^{-\frac{1}{\theta}-1} \right\} dv + \frac{1}{2} (S_1(\tau)^{-\theta} + S_2(\tau)^{-\theta} - 1)^{-1/\theta}.$$

2. The Gumbel copula:

$$p_\tau = \int_{S_2(\tau)}^1 \left\{ \exp \left\{ - [ (-\log S_1(S_2^{-1}(v)))^{\theta+1} + (-\log v)^{\theta+1} ]^{\frac{1}{\theta+1}} \right\} \right. \\ \times \left. \left[ (-\log(S_1(S_2^{-1}(v))))^{\theta+1} + (-\log(v))^{\theta+1} \right]^{-\frac{\theta}{\theta+1}} \frac{(-\log(v))^\theta}{v} \right\} dv \\ + \frac{1}{2} \exp \left\{ - [ (-\log S_1(\tau))^{\theta+1} + (-\log S_2(\tau))^{\theta+1} ]^{\frac{1}{\theta+1}} \right\}.$$

3. The Frank copula:

$$p_\tau = \int_{S_2(\tau)}^1 \frac{e^{-\theta v} \left( e^{-\theta S_1(S_2^{-1}(v))} - 1 \right)}{\left( e^{-\theta} - 1 \right) + \left( e^{-\theta S_1(S_2^{-1}(v))} - 1 \right) \left( e^{-\theta v} - 1 \right)} dv \\ + \frac{1}{2} \left( -\frac{1}{\theta} \log \left( 1 + \frac{(e^{-\theta S_1(\tau)} - 1)(e^{-\theta S_2(\tau)} - 1)}{e^{-\theta} - 1} \right) \right).$$

4. The FGM copula:

$$p_\tau = \int_{S_2(\tau)}^1 \left\{ S_1(S_2^{-1}(v)) + \theta S_1(S_2^{-1}(v)) \left( 1 - S_1(S_2^{-1}(v)) \right) \left( 1 - 2v \right) \right\} dv \\ + \frac{1}{2} (S_1(\tau)S_2(\tau) + \theta S_1(\tau)S_2(\tau)(1 - S_1(\tau))(1 - S_2(\tau))).$$

5. The GB copula:

$$p_\tau = \int_{S_2(\tau)}^1 \left\{ S_1(S_2^{-1}(v)) \left( 1 - \theta \log S_1(S_2^{-1}(v)) \right) v^{-\theta \log S_1(S_2^{-1}(v))} \right\} dv \\ + \frac{1}{2} (S_1(\tau)S_2(\tau) \exp(-\theta \log S_1(\tau) \log S_2(\tau))).$$

### Appendix C. Examples of Using the Shiny Web App

This appendix gives two examples of using the Shiny web app.

**Example A1.** Users set the marginal survival distribution = “Burr III Distribution”,  $c_1 = 1.5$ ,  $k_1 = 3.0$ ,  $c_2 = 1.0$ ,  $k_2 = 1.0$ ,  $\tau = 5.0$ , copula = “FGM”,  $\theta = 0.5$  and the language displayed on the screen “English”, and push the “submit” button, and then they can obtain Kendall’s  $\tau = 0.111$ ,  $p = 0.714$ , and  $p_\tau = 0.719$ . Figure A1 displays the web app in this setting.

Select and Display Results

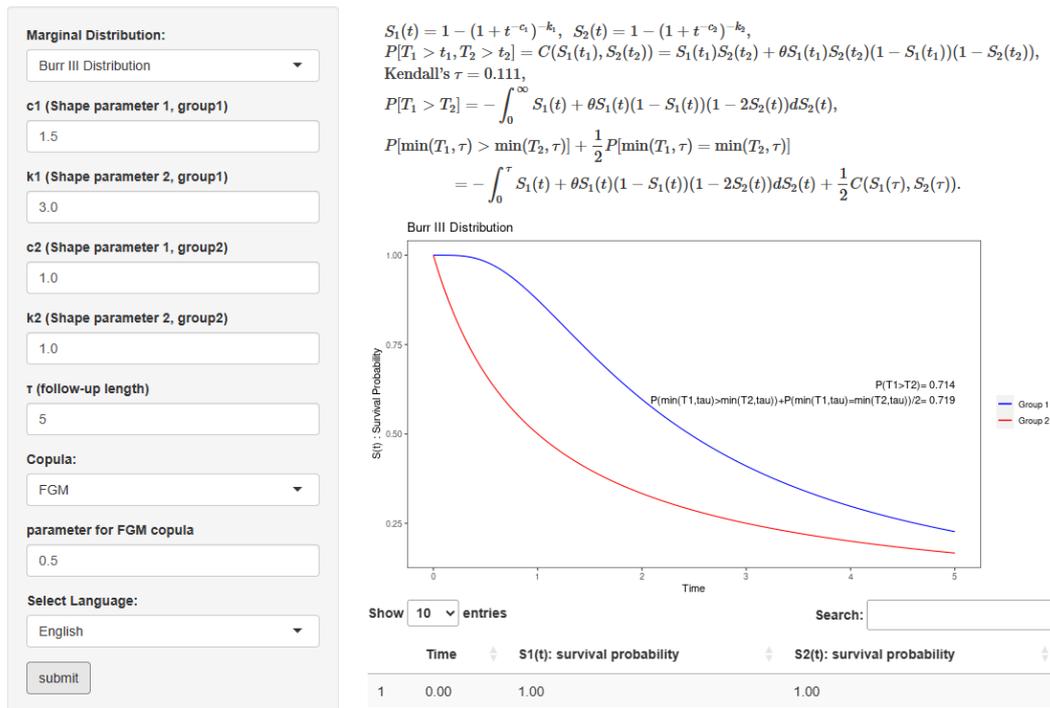


Figure A1. The web app showing the results for computing  $p$  and  $p_{\tau}$  under the Burr III distributions.

**Example A2.** Users set the marginal survival distribution = “Log-normal Distribution”,  $\log(\mu_1) = 0.7$ ,  $\log(\sigma_1^2) = 1.5$ ,  $\log(\mu_2) = 0.3$ ,  $\log(\sigma_2^2) = 2.0$ ,  $\tau = 5.0$ , copula = “Frank”,  $\theta = 2.0$  and the language displayed on the screen “English”, and push the “submit” button, and then they can obtain Kendall’s  $\tau = 0.214$ ,  $p = 0.588$ , and  $p_{\tau} = 0.579$ . Figure A2 displays the web app in this setting.

Select and Display Results

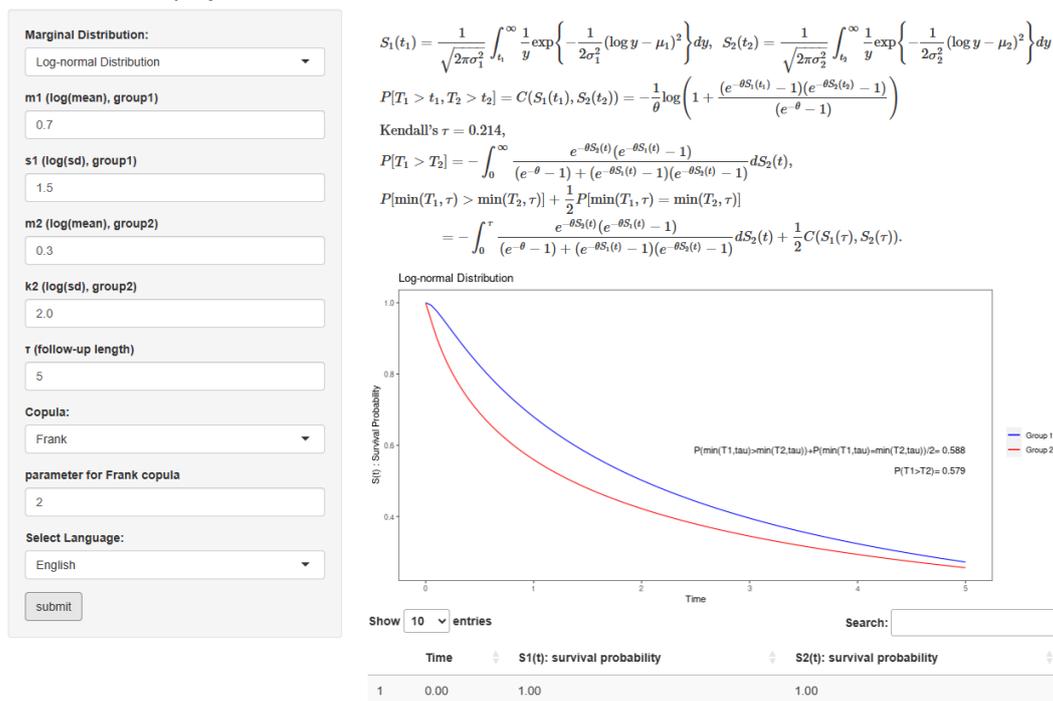


Figure A2. The web app showing the results for computing  $p$  and  $p_{\tau}$  under the log-normal distribution.

## References

1. Efron, B. The two sample problem with censored data. In *Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability*; University of California Press: Berkeley, CA, USA, 1967; Volume 4, pp. 831–853.
2. Rahlfs, V.W.; Zimmermann, H.; Lees, K.R. Effect Size Measures and Their Relationships in Stroke Studies. *Stroke* **2014**, *45*, 627–633. [[CrossRef](#)]
3. Mann, H.B.; Whitney, D.R. On a test of whether one of two random variables is stochastically larger than the other. *Ann. Math. Stat.* **1947**, *18*, 50–60. [[CrossRef](#)]
4. Pocock, S.J.; Ariti, C.A.; Collier, T.J.; Wang, D. The win ratio: A new approach to the analysis of composite endpoints in clinical trials based on clinical priorities. *Eur. Heart J.* **2011**, *33*, 176–182. [[CrossRef](#)]
5. Birnbaum, Z.W. On a use of the Mann-Whitney statistic. In *Proceedings of the Third Berkeley Symposium on Mathematical Statistics and Probability, 1954–1955*; University of California Press: Berkeley, BA, USA, 1956; Volume I, pp. 13–17.
6. Dobler, D.; Pauly, M. Bootstrap- and permutation-based inference for the Mann-Whitney effect for right-censored and tied data. *TEST* **2018**, *27*, 639–658. [[CrossRef](#)]
7. Emura, T.; Hsu, J. Estimation of the Mann-Whitney effect in the two-sample problem under dependent censoring. *Comput. Statist. Data Anal.* **2020**, *150*, 106990. [[CrossRef](#)]
8. Biswas, A.; Chakraborty, S.; Mukherjee, M. On estimation of stress-strength reliability with log-Lindley distribution. *J. Stat. Comput. Simul.* **2021**, *91*, 128–150. [[CrossRef](#)]
9. Rubarth, K.; Sattler, P.; Zimmermann, H.G.; Konietzschke, F. Estimation and Testing of Wilcoxon-Mann-Whitney Effects in Factorial Clustered Data Designs. *Symmetry* **2022**, *14*, 244. [[CrossRef](#)]
10. Hu, J.; Zhuang, Y.; Goldiner, C. Fixed-accuracy confidence interval estimation of  $P(X < Y)$  under a geometric-exponential model. *Jpn. J. Stat. Data Sci.* **2021**, *4*, 1079–1104. [[CrossRef](#)]
11. de la Cruz, R.; Salinas, H.S.; Meza, C. Reliability Estimation for Stress-Strength Model Based on Unit-Half-Normal Distribution. *Symmetry* **2022**, *14*, 837. [[CrossRef](#)]
12. Patil, D.; Naik-Nimbalkar, U.V.; Kale, M.M. Effect of Dependency on the Estimation of  $P[Y > X]$  in Exponential Stress-strength Models. *Austrian J. Stat.* **2022**, *51*, 10–34. [[CrossRef](#)]
13. Nowak, C.P.; Mütze, T.; Konietzschke, F. Group sequential methods for the Mann-Whitney parameter. *Stat. Methods Med. Res.* **2022**, *31*, 2004–2020. [[CrossRef](#)]
14. Singh, B.; Nayal, A.S.; Tyagi, A. Estimation of  $P[Y < Z]$  under Geometric-Lindley model. *Ric. Mat.* **2023**, 1–32. [[CrossRef](#)]
15. Hand, D.J. On Comparing Two Treatments. *Am. Stat.* **1992**, *46*, 190–192. [[CrossRef](#)]
16. Cochran, W.G.; Cox, G.M. *Experimental Designs*, 2nd ed.; John Wiley & Sons Inc.: New York, NY, USA; Chapman & Hall, Ltd.: London, UK, 1957; p. xiv+617.
17. Dobler, D.; Möllenhoff, K. A nonparametric relative treatment effect for direct comparisons of censored paired survival outcomes. *Stat. Med.* **2024**, early view. [[CrossRef](#)] [[PubMed](#)]
18. Fan, Y.; Park, S.S. Sharp bounds on the distribution of treatment effects and their statistical inference. *Econom. Theory* **2010**, *26*, 931–951. [[CrossRef](#)]
19. Fay, M.P.; Brittain, E.H.; Shih, J.H.; Follmann, D.A.; Gabriel, E.E. Causal estimands and confidence intervals associated with Wilcoxon-Mann-Whitney tests in randomized experiments. *Stat. Med.* **2018**, *37*, 2923–2937. [[CrossRef](#)]
20. Emura, T.; Matsui, S.; Rondeau, V. *Survival Analysis with Correlated Endpoints*; SpringerBriefs in Statistics; Springer: Singapore, 2019; p. xvii+118. [[CrossRef](#)]
21. Li, D.; Hu, X.J.; Wang, R. Evaluating association between two event times with observations subject to informative censoring. *J. Amer. Statist. Assoc.* **2023**, *118*, 1282–1294. [[CrossRef](#)]
22. Emura, T.; Pan, C. Parametric likelihood inference and goodness-of-fit for dependently left-truncated data, a copula-based approach. *Statist. Pap.* **2020**, *61*, 479–501. [[CrossRef](#)]
23. Sklar, M. Fonctions de répartition à  $n$  dimensions et leurs marges. *Publ. Inst. Statist. Univ. Paris* **1959**, *8*, 229–231.
24. Nelsen, R.B. *An Introduction to Copulas*, 2nd ed.; Springer Series in Statistics; Springer: New York, NY, USA, 2006; p. xiv+269. [[CrossRef](#)]
25. Geenens, G. (Re-)Reading Sklar (1959); A Personal View on Sklar’s Theorem. *Mathematics* **2024**, *12*, 380. [[CrossRef](#)]
26. Escarela, G.; Carrière, J.F. Fitting competing risks with an assumed copula. *Stat. Methods Med. Res.* **2003**, *12*, 333–349. [[CrossRef](#)] [[PubMed](#)]
27. Petti, D.; Eletti, A.; Marra, G.; Radice, R. Copula link-based additive models for bivariate time-to-event outcomes with general censoring scheme. *Comput. Statist. Data Anal.* **2022**, *175*, 107550. [[CrossRef](#)]
28. 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](#)]
29. Gumbel, E.J. Distributions des valeurs extrêmes en plusieurs dimensions. *Publ. Inst. Statist. Univ. Paris* **1960**, *9*, 171–173.
30. Frank, M.J. On the simultaneous associativity of  $F(x, y)$  and  $x + y - F(x, y)$ . *Aequationes Math.* **1979**, *19*, 194–226. [[CrossRef](#)]
31. Morgenstern, D. Einfache Beispiele zweidimensionaler Verteilungen. *Mitteilungsbl. Math. Statist.* **1956**, *8*, 234–235.
32. Chesneau, C. On the Gumbel-Barnett extended Celebioglu-Cuadras copula. *Jpn. J. Stat. Data Sci.* **2023**, *6*, 759–781. [[CrossRef](#)]

33. Toparkus, A.; Weißbach, R. Testing Truncation Dependence: The Gumbel-Barnett Copula. *arXiv* **2024**, arXiv:2305.19675.
34. Schneider, S.; dos Reis, R.C.P.; Gottselig, M.M.F.; Fisch, P.; Knauth, D.R.; Vigo, A. Clayton copula for survival data with dependent censoring: An application to a tuberculosis treatment adherence data. *Stat. Med.* **2023**, *42*, 4057–4081. [[CrossRef](#)]
35. Sun, T.; Ding, Y. Copula-based semiparametric regression method for bivariate data under general interval censoring. *Biostatistics* **2021**, *22*, 315–330. [[CrossRef](#)]
36. Moradian, H.; Larocque, D.; Bellavance, F. Survival forests for data with dependent censoring. *Stat. Methods Med. Res.* **2019**, *28*, 445–461. [[CrossRef](#)] [[PubMed](#)]
37. Farzana, W.; Basree, M.M.; Diawara, N.; Shboul, Z.A.; Dubey, S.; Lockhart, M.M.; Hamza, M.; Palmer, J.D.; Iftekharruddin, K.M. Prediction of Rapid Early Progression and Survival Risk with Pre-Radiation MRI in WHO Grade 4 Glioma Patients. *Cancers* **2023**, *15*, 4636. [[CrossRef](#)] [[PubMed](#)]
38. Chen, Y. Semiparametric marginal regression analysis for dependent competing risks under an assumed copula. *J. R. Stat. Soc. Ser. B Stat. Methodol.* **2010**, *72*, 235–251. [[CrossRef](#)]
39. Shih, J.; Emura, T. Likelihood-based inference for bivariate latent failure time models with competing risks under the generalized FGM copula. *Comput. Statist.* **2018**, *33*, 1293–1323. [[CrossRef](#)]
40. Kawakami, R.; Michimae, H.; Lin, Y. Assessing the numerical integration of dynamic prediction formulas using the exact expressions under the joint frailty-copula model. *Jpn. J. Stat. Data Sci.* **2021**, *4*, 1293–1321. [[CrossRef](#)]
41. Shih, J.; Konno, Y.; Chang, Y.; Emura, T. Estimation of a common mean vector in bivariate meta-analysis under the FGM copula. *Statistics* **2019**, *53*, 673–695. [[CrossRef](#)]
42. Shih, J.; Konno, Y.; Chang, Y.; Emura, T. Copula-Based Estimation Methods for a Common Mean Vector for Bivariate Meta-Analyses. *Symmetry* **2022**, *14*, 186. [[CrossRef](#)]
43. Dobler, D.; Pauly, M. Factorial analyses of treatment effects under independent right-censoring. *Stat. Methods Med. Res.* **2020**, *29*, 325–343. [[CrossRef](#)] [[PubMed](#)]
44. Emura, T.; Ditzhaus, M.; Dobler, D.; Murotani, K. Factorial survival analysis for treatment effects under dependent censoring. *Stat. Methods Med. Res.* **2024**, *33*, 61–79. [[CrossRef](#)]
45. Moore, D.F. *Applied Survival Analysis Using R*; Springer: Berlin/Heidelberg, Germany, 2016; Volume 473.
46. Greenland, S.; Fay, M.P.; Brittain, E.H.; Shih, J.H.; Follmann, D.A.; Gabriel, E.E.; Robins, J.M. On causal inferences for personalized medicine: How hidden causal assumptions led to erroneous causal claims about the *D*-value. *Am. Statist.* **2020**, *74*, 243–248. [[CrossRef](#)]
47. Domma, F.; Giordano, S. A copula-based approach to account for dependence in stress-strength models. *Statist. Pap.* **2013**, *54*, 807–826. [[CrossRef](#)]
48. Gao, J.; An, Z.; Liu, B. A dependent stress-strength interference model based on mixed copula function. *J. Mech. Sci. Technol.* **2016**, *30*, 4443–4446. [[CrossRef](#)]
49. de Andrade, B.B.; do Nascimento, A.R.; Rathie, P.N. Parametric and nonparametric inference for the reliability of copula-based stress-strength models. *Qual. Reliab. Eng. Int.* **2020**, *36*, 2249–2267. [[CrossRef](#)]
50. Rathie, P.N.; de Sena Monteiro Ozelim, L.C.; de Andrade, B.B. Portfolio Management of Copula-Dependent Assets Based on  $P(Y < X)$  Reliability Models: Revisiting Frank Copula and Dagum Distributions. *Stats* **2021**, *4*, 1027–1050. [[CrossRef](#)]
51. James, A.; Chandra, N.; Sebastian, N. Stress-strength reliability estimation for bivariate copula function with rayleigh marginals. *Int. J. Syst. Assur. Eng. Manag.* **2023**, *14*, 196–215. [[CrossRef](#)]
52. Shang, L.; Yan, Z. Reliability estimation stress-strength dependent model based on copula function using ranked set sampling. *J. Radiat. Res. Appl. Sci.* **2024**, *17*, 100811. [[CrossRef](#)]
53. Lima, R.K.; Quintino, F.S.; da Fonseca, T.A.; de Sena Monteiro Ozelim, L.C.; Rathie, P.N.; Saulo, H. Assessing the Impact of Copula Selection on Reliability Measures of Type  $P(X < Y)$  with Generalized Extreme Value Marginals. *Modelling* **2024**, *5*, 180–200. [[CrossRef](#)]
54. Patton, A.J. Modelling asymmetric exchange rate dependence. *Internat. Econom. Rev.* **2006**, *47*, 527–556. [[CrossRef](#)]
55. Almeida, C.; Czado, C. Efficient Bayesian inference for stochastic time-varying copula models. *Comput. Statist. Data Anal.* **2012**, *56*, 1511–1527. [[CrossRef](#)]
56. Yin, Y.; Cai, Z.; Zhou, X. Using secondary outcome to sharpen bounds for treatment harm rate in characterizing heterogeneity. *Biom. J.* **2018**, *60*, 879–892. [[CrossRef](#)]
57. Susam, S.O. A multi-parameter Generalized Farlie-Gumbel-Morgenstern bivariate copula family via Bernstein polynomial. *Hacet. J. Math. Stat.* **2022**, *51*, 618–631. [[CrossRef](#)]
58. Deresa, N.W.; Van Keilegom, I. A multivariate normal regression model for survival data subject to different types of dependent censoring. *Comput. Statist. Data Anal.* **2020**, *144*, 106879. [[CrossRef](#)]
59. Jo, J.H.; Gao, Z.; Jung, I.; Song, S.Y.; Ridder, G.; Moon, H.R. Copula graphic estimation of the survival function with dependent censoring and its application to analysis of pancreatic cancer clinical trial. *Stat. Methods Med. Res.* **2023**, *32*, 944–962. [[CrossRef](#)] [[PubMed](#)]

- 
60. Emura, T.; Chen, Y. *Analysis of Survival Data with Dependent Censoring: Copula-Based Approaches*; Springer: Berlin/Heidelberg, Germany, 2018; Volume 450.
  61. Brunner, E.; Puri, M.L. Nonparametric methods in factorial designs. *Statist. Pap.* **2001**, *42*, 1–52. [[CrossRef](#)]

**Disclaimer/Publisher’s Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.