



Weighted U-statistics for likelihood-ratio ordering of bivariate data

Sangita Kulathinal¹ · Isha Dewan²

Received: 13 September 2021 / Revised: 27 January 2022 / Accepted: 21 May 2022 /
Published online: 2 July 2022
© The Author(s) 2022

Abstract

Characterisation of marginal distribution and density functions is of interest where data on a pair of random variables (X, Y) are observed. Stochastic orderings between (X, Y) have been studied in statistics and economics. Likelihood-ratio ordering is useful in understanding the behaviour of the random variables. In this article, tests based on U-statistics are proposed to test for equality of marginal density functions against the alternative of likelihood-ratio ordered when (X, Y) are dependent. The tests can be used when the data are either completely observed or subjected to independent univariate right censoring. The asymptotic variances of these tests are complicated and hence, are estimated using jackknife variance estimators. Validity of the jackknife variance estimators in statistical inference based on the proposed tests is demonstrated using simulation studies. The test for uncensored setting has desired size and good power for small sample. The performance of the tests for censored case depends on the sample size, proportion of censoring and the measure of dependence between X and Y . The tests are illustrated on three real data sets chosen in order to bring out various aspects of the tests.

Keywords U-statistics · Censored data · Bivariate distributions · Likelihood-ratio ordering · Kaplan-Meier estimator · Inverse probability of censoring weights

Sangita Kulathinal's work was partly supported by the grant for research mobility abroad of the Academy of Finland (No. 325990). Isha Dewan thanks the Science and Engineering Research Board (SERB), Department of Science and Technology for its research grant.

✉ Sangita Kulathinal
sangita.kulathinal@helsinki.fi

¹ Department of Mathematics and Statistics, University of Helsinki, Helsinki, Finland

² Indian Statistical Institute, New Delhi, India

1 Introduction

Bivariate data arise in different fields such as biomedicine, epidemiology, insurance, reliability, survival analysis, and social science. The pairing of the observed times occur either naturally or artificially in many studies. Common examples include an event observed in a matched pair or twins, progression of a disease in both eyes, damage in right and left joints, mortality of insured couples, events such as shedding of cytomegalovirus (CMV) in the urine and blood, colonization of mycobacterium avium complex (MAC) in the sputum and stool, or CMV and MAC in AIDS patients. For illustrations of such data in practice we refer to (Lin and Ying 1993; Betensky and Finkelstein 1999; Lu and Burke 2008; Savignoni et al. 2014; Shigemizu et al 2017; Peres et al. 2020). A critical review of statistical models is provided by Prentice and Kalbfleisch (2003).

Families of bivariate distributions used to model paired data are discussed by Lai et al. (2006). Often the interest lies in the marginal distributions and comparisons accounting for the dependency between such pairs (Luciano et al. 2008; Nair et al. 2018; Peres et al. 2020). Any bivariate distribution function can be represented in terms of the marginal ones and a copula function (Sklar's theorem, Nelsen 1999). Hence, any understanding of the behaviour of marginals would help in building a realistic bivariate model.

Stochastic orderings between a pair of random variables have been of interest in statistics and economics. These orderings are discussed in detail in Shaked and Shanthikumar (2007). A random variable X with distribution function F is said to be stochastically greater than a random variable Y with distribution function G , denoted as $Y \stackrel{\text{st}}{<} X$, if $F(t) \leq G(t)$ for all t . Sometimes a stronger condition of likelihood-ratio (LR) ordering may hold. Suppose f and g are densities corresponding to F and G . We say that X is greater than Y according to likelihood-ratio ordering, $Y \stackrel{\text{lr}}{<} X$, if the ratio of their respective densities $f(t)/g(t)$ is nondecreasing in t . LR ordering is preserved under increasing transformations, i.e., $Y \stackrel{\text{lr}}{<} X$, then $\phi(Y) \stackrel{\text{lr}}{<} \phi(X)$ for all real valued increasing functions ϕ . Many families of random variables, e.g., one-parameter exponential family of distributions, are likelihood-ratio ordered according to some parameter of interest (Table 2.1, page 102, Belzunce et al. 2016). LR ordering has been referred to as uniform stochastic ordering (Keilson and Sumita 1982) and also as density-ratio ordering (Beare and Moon 2015).

Ross (1983) and Shanthikumar et al. (1991) have discussed applications of LR ordering in stochastic scheduling, closed queuing network and reliability. Its applications in portfolio choice, crop insurance, mechanism design and auction theory are discussed in Roosen and Hennessy (2004). In empirical finance—pricing kernel is the ratio of risk neutral and physical densities for the payoff distribution of a market index at a date in future. In classical finance this ratio is assumed to be monotonic.

The problem of testing LR ordering has been addressed in the literature for discrete as well as continuous distributions based on independent samples from F and G (see for example, Dykstra et al. 1995; Xiong et al. 2002; Roosen and Hennessy 2004; Carolan and Tebbs 2005; Beare and Moon 2015; Yu et al 2017; Westling et al.

2019). This article focuses on testing for LR ordering based on independent samples from the bivariate distribution of (X, Y) with marginals f and g .

In the case of bivariate event times either or both times may be subjected to censoring. The data may be censored due to study design, loss to follow-up, withdrawal etc. A number of authors have studied the analysis of bivariate event time data when only univariate independent right censoring is present. Lin and Ying (1993) proposed a simple nonparametric estimator of bivariate survival function under univariate censoring. Since then the literature on bivariate event times has geared to modifications in the estimator of the joint survival function under various censoring schemes (Geerdens et al. 2016; Jinnah and Zhao 2017; Huang et al. 2018). Hutson (2016) have used a constrained maximum likelihood density based approach to study a bivariate estimator of the joint survival function where marginals are Kaplan-Meier estimators. Marginal modelling of bivariate data using, e.g., frailty has also been of interest (Rondeau et al. 2012; Begun and Yashin 2018). Copula-based modelling has got attention during the last decades (Luciano et al. 2008; Diao and Cook 2014; Peres et al. 2020). More research on characterisation of marginal densities for bivariate data, complete or censored, is needed.

U-statistics are very widely used and elegant tools for estimation and testing (Hoeffding 1948; Lee 2020). We develop U-statistics based tests for testing equality of the marginal density functions against LR ordering for uncensored as well as censored bivariate data. The censoring scheme is assumed to be independent univariate censoring described in Lin and Ying (1993).

The article is organised as follows. In Sect. 2, we propose a U-statistic for testing equality of the marginal density functions against the LR ordering for complete data. We extend this test to independent univariate censoring case in Sect. 3 and propose weighted U-statistics. We derive asymptotic distributions for the proposed U-statistics. For the weighted U-statistics, we sketch proofs of asymptotic normality and refer to Datta et al. (2010) for more details. We employ jackknife method to estimate variances of weighted U-statistics. Extensive simulation studies are carried out to study and compare the performance of the tests in Sect. 4. Necessary mathematical details are included in appendices A-E. Tests are applied to three real data sets in Sect. 5. The article ends with a detailed discussion.

2 Test for likelihood-ratio ordering: no censoring

Suppose $Z_i = (X_i, Y_i)$, $i = 1, \dots, n$ are independent and identically distributed bivariate samples from joint probability density and distribution functions $w(x, y)$ and $W(x, y)$, and with marginal density and survival functions $(f(x), \bar{F}(x))$ and $(g(y), \bar{G}(y))$, respectively. Without loss of generality, we consider positive-valued random variables. Consider a hypothesis testing problem of the equality of densities against the likelihood-ratio ordering, i.e., $H_{0L} : f(t)/g(t) = 1, \forall t$ vs. $H_{1L} : f(t)/g(t) \uparrow t$. Under H_{1L} , for $s < t$

$$\frac{f(s)}{g(s)} \leq \frac{f(t)}{g(t)}, \quad \text{i.e., } f(t)g(s) - f(s)g(t) \geq 0, \quad (1)$$

and strict inequality holds for some $s < t$. Consider the following functional by integrating (1) over $s < t$

$$\int_{s < t} \left[f(t)g(s) - f(s)g(t) \right] ds dt. \quad (2)$$

Note that (2) can be written as

$$\Delta_L(F, G) = P(Y_2 < X_1) - P(Y_2 > X_1) = 2P(Y_2 < X_1) - 1, \quad (3)$$

where (X_1, Y_1) and (X_2, Y_2) are two independent copies of bivariate random variables with joint density $w(x, y)$. The functional $\Delta_L(F, G)$ is zero under H_{0L} and positive under H_{1L} .

Consider a symmetric kernel $\phi(Z_i, Z_j)$ of degree 2 which is an unbiased estimator of the functional $\Delta(F, G) = \Delta_L(F, G) + 1$.

$$\phi(Z_i, Z_j) = \phi(X_i, Y_i, X_j, Y_j) = \begin{cases} 1, & (Y_j < X_i) \text{ or } (Y_i < X_j) \\ 0, & \text{otherwise.} \end{cases} \quad (4)$$

Next, we define the corresponding U-statistic:

$$U_L = \frac{1}{\binom{n}{2}} \sum_{1 \leq i < j \leq n} \phi(Z_i, Z_j). \quad (5)$$

Note that $E(U_L)$ is 1 under H_{0L} and greater than 1 under H_{1L} .

The following observations can be made immediately. The statistic U_L can be seen as a sign statistic as it counts the proportion of positive differences in $\binom{n}{2}$ symmetric pairs of differences $(X_i - Y_j)$ and $(X_j - Y_i)$. The actual magnitude of the differences is not under consideration. The statistic U_L counts the number of times an X observation exceeds a Y observation excluding the difference between X and Y from the same pair. In this sense it can be seen as a version of the Mann-Whitney statistic.

For the ease of computation, we can express the above U-statistic as a function of ranks in the absence of ties. Let $X_{(1)} < \dots < X_{(n)}$ be ordered X observations and let R_i be the rank of $X_{(i)}$ in the combined sample of $2n$ observations, (X_i, Y_i) , $i = 1, \dots, n$. The U-statistics can be expressed as:

$$\begin{aligned} U_L &= \frac{1}{\binom{n}{2}} \sum_{i=1}^n \left[R_i - i - I(Y_{(i)} < X_{(i)}) \right], \\ &= \frac{1}{\binom{n}{2}} \left[\sum_{i=1}^n R_i - n(n+1)/2 - n_{I(Y < X)} \right], \end{aligned} \quad (6)$$

where $Y_{(i)}$ indicates the Y observation corresponding to $X_{(i)}$ and $n_{I(Y < X)}$ is the number of pairs $(X_{(i)}, Y_{(i)})$ where $Y_{(i)} < X_{(i)}$. Here and in the sequel, $I(\cdot)$ denotes the indicator function.

In order to obtain the asymptotic variance of the U-statistic, we first note that for fixed $z_i = (x_i, y_i)$ the expectation of the kernel (4) w.r.t. $w(x, y)$ is

$$\phi_1(z_i) = E(\phi(Z_i, Z_j)|z_i) = G(x_i) + \bar{F}(y_i).$$

Note that $E(U_L) = E(\phi(Z_i, Z_j)) = E(\phi_1(Z_i))$ and the variance of $\phi_1(Z_i)$, $\sigma_1^2 = E(\phi_1^2(Z_i)) - (E(\phi_1(Z_i)))^2$ where the expectations w.r.t. $w(x, y)$ are

$$\begin{aligned} E(\phi_1(Z_i)) &= \int_0^\infty G(x)f(x)dx + \int_0^\infty \bar{F}(y)g(y)dy, \\ E(\phi_1^2(Z_i)) &= \int_0^\infty G^2(x)f(x)dx + \int_0^\infty \bar{F}^2(y)g(y)dy \\ &\quad + 2 \int_0^\infty \int_0^\infty G(x)\bar{F}(y)w(x, y)dydx. \end{aligned} \tag{7}$$

The following theorem gives the asymptotic normality of the statistic U_L .

Theorem 1 *If $E(\phi^2(Z_i, Z_j)) < \infty$ and $\sigma_1^2 > 0$ then $\sqrt{n}(U_L - E(U_L))$ converges in distribution to $N(0, 4\sigma_1^2)$ as $n \rightarrow \infty$. Under the null hypothesis $E(U_L) = 1$ and $\sigma_1^2 = 2P(X_2 < X_1, Y_3 > Y_1) - 1/3$.*

Proof The asymptotic normality of the U-statistic follows from Hoeffding’s decomposition (Hoeffding 1948; Lee 2020). Under the null hypothesis $f(\cdot) = g(\cdot)$, (7) simplifies and we obtain

$$E(\phi(Z_i, Z_j)) = E(\phi_1(Z_i)) = 1, \tag{8}$$

and the variance of $\phi_1(X_i, Y_i)$ is

$$\begin{aligned} E(\phi_1^2(Z_i)) &= 2/3 + 2 \int_0^\infty \int_0^\infty F(x)\bar{F}(y)w(x, y)dydx, \text{ and} \\ \sigma_1^2 &= 2P(X_2 < X_1, Y_3 > Y_1) - \frac{1}{3}. \end{aligned} \tag{9}$$

□

It is obvious that σ_1^2 depends on the joint distribution of (X, Y) . Under the null hypothesis, one can estimate (9) by the U-statistic associated with the kernel $I(X_j < X_i, Y_k > Y_i)$ as:

$$\hat{\sigma}_1^2 = \frac{2}{nP_3} \sum_{i \neq j \neq k=1}^n I(X_j < X_i, Y_k > Y_i) - \frac{1}{3}. \tag{10}$$

Large values of the U-statistic indicate that the observed sample is an outlier under the null hypothesis of equality of the marginal densities. The statistic U_L can also be used to test H_{0L} vs H_{1L} when X and Y are independent. When X and Y are

independent, the expectation $E(U_L)$ is 1, and the asymptotic variance under H_{0L} simplifies to 4/6. It provides a very simple alternative distribution-free test procedure compared to other existing tests in the literature mentioned in the Introduction.

3 Tests for likelihood-ratio ordering: independent univariate censoring

Under the censoring setting described in Lin and Ying (1993), the observed data are $(T_i, \delta_i^x, U_i, \delta_i^y)$, $i = 1, \dots, n$ where $T_i = X_i \wedge C_i$, $\delta_i^x = I(X_i \leq C_i)$, $U_i = Y_i \wedge C_i$ and $\delta_i^y = I(Y_i \leq C_i)$. Here $a \wedge b = \min(a, b)$ and $a \vee b = \max(a, b)$. The censoring time C , with survival function $S_c(\cdot)$, is assumed to be independent of (X, Y) . We extend the proposed U-statistics to this censoring scheme. Throughout this article, proportion of censoring refers to $\#(\delta_i^x = 0 \text{ or } \delta_i^y = 0)/n$ or equivalently, $1 - \#(\delta_i^x = 1 \text{ and } \delta_i^y = 1)/n$.

3.1 Direct extension of (5)

Consider the following symmetric kernel which is a direct extension of the kernel $\phi(Z_i, Z_j)$, defined in (4), to the censoring scheme described above.

$$\phi_{0c}(\tilde{Z}_i, \tilde{Z}_j) = \begin{cases} \frac{1}{2}, & (\delta_j^y = 1, U_j < T_i) \text{ or } (\delta_i^y = 1, U_i < T_j) \\ -\frac{1}{2}, & (\delta_i^x = 1, U_j > T_i) \text{ or } (\delta_j^x = 1, U_i > T_j) \\ 0, & \text{otherwise,} \end{cases} \quad (11)$$

where $\tilde{Z}_i = (T_i, \delta_i^x, U_i, \delta_i^y)$. Note that $(\delta_j^y = 1, U_j < T_i)$ means that $Y_j < C_j$ and $Y_j < X_i \wedge C_i$. This ensures that $Y_j < X_i$ and hence, the kernel is given a positive value. For $(\delta_i^x = 1, U_j > T_i)$, $X_j < C_j$ and $X_j < Y_i \wedge C_i$. This implies that $X_j < Y_i$ and hence, the kernel is given the same value with negative sign. Other cases can be interpreted similarly.

Define a U-statistic in the presence of univariate independent censoring as

$$U_{0c} = \frac{1}{\binom{n}{2}} \sum_{1 \leq i < j \leq n} \phi_{0c}(\tilde{Z}_i, \tilde{Z}_j). \quad (12)$$

The expectation of (12) is:

$$\begin{aligned} E(U_{0c}) &= E(\phi_{0c}(\tilde{Z}_i, \tilde{Z}_j)) = \left[\int_0^\infty S_c^2(y) \bar{F}(y) g(y) dy - \int_0^\infty S_c^2(x) \bar{G}(x) f(x) dx \right], \\ &= 0, \quad \text{when } f(\cdot) = g(\cdot) \\ &> 0, \quad \text{under } H_{1L}. \end{aligned} \quad (13)$$

The last inequality can be obtained by multiplying (1) by $S_c^2(s)$ and then integrating both sides as in (2). In the absence of censoring, $S_c(t) = 1, \forall t$ and hence, $E(\phi_{0c}(\tilde{Z}_i, \tilde{Z}_j)) = \Delta_L(F, G)$. The asymptotic distribution of U_{0c} follows from the standard theory of U-statistics (cf. Appendix A).

3.2 Weighted U-statistics

We modify the kernel $\phi(Z_i, Z_j)$ defined in (4) by using the inverse probability of censoring as weights in two different ways.

I. Inverse probability of censoring weighted symmetric kernel We extend (5) by using the inverse probability of censoring weighted U-statistic as in Datta et al. (2010). We first define the observed data to resemble their setting. Let $L_i^* = X_i \vee Y_i, L_i = L_i^* \wedge C_i$ and $\delta_i = I(L_i^* \leq C_i)$. The observed data are $(L_i, \delta_i, V_i = Z_i \delta_i), i = 1, \dots, n$. Note that $\delta_i = 1$ if and only if both $X_i \leq C_i$ and $Y_i \leq C_i$, i.e., $\delta_i^x = \delta_i^y = 1$. Hence, $Z_i = (X_i, Y_i)$ is observed. The probability that Z_i is observed is $P(X_i \leq C_i, Y_i \leq C_i) = S_c(X_i \vee Y_i -)$. An estimator analogous to (5) can be obtained as a weighted average of $\phi(Z_i, Z_j)$ when both Z_i and Z_j are observed. The weight for each pair (i, j) is the inverse of the probability that they are observed. This probability is $S_c(X_i \vee Y_i -)S_c(X_j \vee Y_j -)$. The inverse probability weighted U-statistic corresponding to (5) is defined as follows.

$$\begin{aligned}
 U_{2c} &= \frac{1}{\binom{n}{2}} \sum_{1 \leq i < j \leq n} \phi(V_i, V_j) \left[\frac{\delta_i^x \delta_i^y \delta_j^x \delta_j^y}{S_c(L_i -) S_c(L_j -)} \right] \\
 &= \frac{1}{\binom{n}{2}} \sum_{1 \leq i < j \leq n} \phi(Z_i, Z_j) \left[\frac{\delta_i^x \delta_i^y \delta_j^x \delta_j^y}{S_c(X_i \vee Y_i -) S_c(X_j \vee Y_j -)} \right] \\
 &= \frac{1}{\binom{n}{2}} \sum_{1 \leq i < j \leq n} (I(Y_j < X_i) + I(Y_i < X_j)) \left[\frac{\delta_i^x \delta_i^y \delta_j^x \delta_j^y}{S_c(X_i \vee Y_i -) S_c(X_j \vee Y_j -)} \right] \\
 &= \frac{1}{\binom{n}{2}} \sum_{1 \leq i < j \leq n} (I(U_j < T_i) + I(U_i < T_j)) \left[\frac{\delta_i^x \delta_i^y \delta_j^x \delta_j^y}{S_c(T_i \vee U_i -) S_c(T_j \vee U_j -)} \right].
 \end{aligned}
 \tag{14}$$

The second equality follows because $V_i = Z_i$ when $\delta_i = 1$. The U-statistic (14) is mean preserving. i.e., $E(U_{2c}) = E(U_L) = \Delta(F, G) = 2P(Y < X)$. As pointed out in Datta et al. (2010), (14) is a standard U-statistic if $S_c(\cdot)$ is known and its asymptotic distribution can be obtained using the standard theory of U-statistics (Appendix B).

In practice, the censoring survival function is unknown and can be estimated by the Kaplan-Meier estimator (Andersen et al. 1993). The censoring time C_i is observed when either δ_i^x and/or δ_i^y are zero and it is censored by $X_i \vee Y_i$ when both δ_i^x and δ_i^y are 1. Let $\hat{S}_c(\cdot)$ be the Kaplan-Meier estimator of the survival function of censoring time C . An estimator \hat{U}_{2c} of U_{2c} is obtained by replacing the survival function $S_c(t)$

of the censoring time by its Kaplan-Meier estimator $\hat{S}_c(t)$.

$$\hat{U}_{2c} = \frac{1}{\binom{n}{2}} \sum_{1 \leq i < j \leq n} (I(U_j < T_i) + I(U_i < T_j)) \left[\frac{\delta_i^x \delta_i^y \delta_j^x \delta_j^y}{\hat{S}_c(T_i \vee U_i-) \hat{S}_c(T_j \vee U_j-)} \right]. \tag{15}$$

Note that (15) is not a standard U-statistic since the kernel depends on all data through the Kaplan-Meier estimator. Its asymptotic distribution can be shown to be normal with asymptotic variance being complicated, as shown by Datta et al. (2010) for the univariate case. A sketch of the proof is given in (Appendix B).

II. Inverse probability of censoring weighted asymmetric kernel Only pairs of uncensored bivariate data weighted by their joint probability of being observed, i.e., only pairs (X_i, Y_i) and (X_j, Y_j) with $\delta_i^x = \delta_i^y = \delta_j^x = \delta_j^y = 1$ contribute to the U-statistic (14). Note that the kernel (4) is defined using $I(Y_j < X_i)$, $i \neq j$ and X_i is observed with probability $P(X_i \leq C_i) = S_c(X_i-)$, and Y_j is observed with probability $P(Y_j \leq C_j) = S_c(Y_j-)$. The probability that both (X_i, Y_j) are observed is $P(X_i \leq C_i, Y_j \leq C_j) = S_c(X_i-)S_c(Y_j-)$. An estimator analogous to (5) can be obtained as a weighted average of $I(Y_j < X_i)$ when both X_i and Y_j are observed, ($\delta_i^x = 1$) and ($\delta_j^y = 1$), and $I(Y_i < X_j)$ when both X_j and Y_i are observed, ($\delta_j^y = 1$) and ($\delta_i^x = 1$). The weights are the inverse of the probabilities that they are observed. These probabilities are $S_c(X_i-)S_c(Y_j-)$ and $S_c(Y_i-)S_c(X_j-)$, respectively. This is equivalent to first multiplying each term of the kernel (4) by the inverse probability of censoring, and then define a symmetric kernel as:

$$\begin{aligned} \phi_{1c}(\tilde{Z}_i, \tilde{Z}_j) &= I(Y_j < X_i) \frac{\delta_i^x \delta_j^y}{S_c(X_i-)S_c(Y_j-)} + I(Y_i < X_j) \frac{\delta_i^y \delta_j^x}{S_c(Y_i-)S_c(X_j-)} \\ &= I(U_j < T_i) \frac{\delta_i^x \delta_j^y}{S_c(T_i-)S_c(U_j-)} + I(U_i < T_j) \frac{\delta_i^y \delta_j^x}{S_c(U_i-)S_c(T_j-)}. \end{aligned} \tag{16}$$

The modified U-statistic can then be given as

$$U_{1c} = \frac{1}{\binom{n}{2}} \sum_{1 \leq i < j \leq n} \phi_{1c}(\tilde{Z}_i, \tilde{Z}_j). \tag{17}$$

The above U-statistic has contribution from observations i and j provided $\delta_i^x = \delta_j^y = 1$ and hence, δ_i^y and δ_j^x may be either 0 or 1. The expectation $E(U_{1c}) = \Delta(F, G)$ as can be seen from the following derivation.

Consider the expectation of first term in the sum with respect to the joint distribution of $(T_i, \delta_i, U_j, \delta_j)$. Note that (T_i, δ_i) and (U_j, δ_j) are independent and the densities of

$(T_i, \delta_i^x = 1)$ and $(U_j, \delta_j^y = 1)$ are $S_c(t)f(t)$ and $S_c(t)g(t)$, respectively.

$$\begin{aligned} & E\left(I(U_j < T_i) \frac{\delta_i^x \delta_j^y}{S_c(T_i-)S_c(U_j-)}\right) \\ &= \int_0^\infty \int_0^x \frac{1}{S_c(x-)S_c(y-)} S_c(y-)g(y)S_c(x-)f(x)dydx \\ &= \int_0^\infty \int_0^x g(y)f(x)dydx = P(Y_j < X_i). \end{aligned}$$

It is easy to see that U_{1c} is a U-statistic and we can apply the theory of U-statistic to obtain the asymptotic distribution (cf. Appendix C). An estimator \widehat{U}_{1c} of U_{1c} , as in the previous section, is obtained by replacing the survival function $S_c(t)$ of the censoring time by the Kaplan-Meier estimator $\widehat{S}_c(t)$.

$$\widehat{U}_{1c} = \frac{1}{\binom{n}{2}} \sum_{1 \leq i < j \leq n} \widehat{\phi}_{1c}(\widetilde{Z}_i, \widetilde{Z}_j). \tag{18}$$

The statistic (18) is not a standard U-statistic. We outline the proof of its asymptotic normality in Appendix C.

Large values of the standardised U-statistics proposed in this section indicate that the observed sample is an outlier under the null hypothesis of equality of the marginal densities. Both (15) and (18) are plug-in estimators and hence, in general, biased. But, asymptotically both are unbiased. The asymptotic distributions depend on the unknown joint distribution of (X, Y) as well as the censoring distribution. For the purpose of hypothesis testing, a good estimator for the asymptotic variance would be required. We investigate the empirical densities and variance estimators in the weighted U-statistics (15) and (18) in the next section.

Expectations of tests proposed in Sects. 2 and 3 under the LR ordering H_{1L} are larger than the respective expectations under the null H_{0L} . In each case, large values of the statistic suggest that the observed data may be an outlier under H_{0L} against H_{1L} . Hence, the tests are consistent against the alternative of LR ordering according to the theory behind consistency of U-statistics (Lehmann 1951).

4 Simulation

Monte Carlo studies are carried out to evaluate and compare the performance of the proposed test statistics $U_L, U_{0c}, \widehat{U}_{1c}$ and \widehat{U}_{2c} . All simulations and analyses are carried out in a statistical computing environment R (R core team 2018). For computation of Kaplan-Meier estimates package *survival* (Therneau 2012) and for density estimation package *mudens* (Herrick et al. 2018) from R are used.

We simulate data from Gumbel’s family of bivariate exponential distributions (Gumbel 1960) with the survival and density functions as

$$\widetilde{W}(x, y) = e^{-(\lambda_1 x + \lambda_2 y + \delta \lambda_1 \lambda_2 xy)}, \quad x, y > 0, \lambda_1, \lambda_2 > 0, 0 \leq \delta \leq 1,$$

$$w(x, y) = W(x, y) \{(1 - \delta)\lambda_1\lambda_2 + \delta\lambda_1^2\lambda_2x + \delta\lambda_1\lambda_2^2y + \delta^2\lambda_1^2\lambda_2^2xy\}. \quad (19)$$

The ratio of the marginal densities $f(t)/g(t)$ is proportional to $e^{-(\lambda_1 - \lambda_2)t}$. This ratio is increasing when $\lambda_1 < \lambda_2$. Thus we have the likelihood-ratio ordering between X and Y when $\lambda_1 < \lambda_2$. We allow the marginal exponential hazards to be $\lambda_1 = 1$ and $\lambda_2 = (1, 1.2, 1.5, 2)$. This family models negative association for $0 \leq \delta \leq 1$, wherein $\delta = 0$ corresponds to independence. Simulation algorithm is described in Appendix D.

4.1 No censoring: performance of U_L

We generate 2000 random samples from the bivariate exponential distribution with varying degree of dependence and samples sizes to study empirical size and power of the test in the absence of censoring. The samples of sizes $n = (25, 50, 100, 200, 500)$ are generated with $\lambda_2 = 1$ to evaluate the empirical size and with $\lambda_2 = (1.2, 1.5, 2)$ to evaluate the power of the tests.

The variance σ_1^2 under the null hypothesis, $\lambda_1 = \lambda_2 = 1$, is evaluated analytically using (9) for the bivariate exponential distribution. The empirical size of the test (5) is obtained by simulating $k = 1, \dots, K$ ($= 2000$), sets of bivariate samples $(X_i^{(k)}, Y_i^{(k)})$, $i = 1, \dots, n$ under the null hypothesis for $\lambda_1 = \lambda_2 = 1$ and $\delta = (0, 0.2, 0.5, 1)$ for different choices of n , and computing the standardised U-statistic, $z_k = \sqrt{n}(U_{1L}^{*(k)} - 1)/\sqrt{(4 * \sigma_1^2)}$. The empirical size is then computed as

$$\alpha_{1L} = \frac{1}{K} \sum_{k=1}^K I(z_k > z_\alpha),$$

where z_α is the $(1 - \alpha)$ quantile of the standard normal distribution. Recall that for $\alpha = 5\%$, $z_\alpha = 1.64$. The power is obtained by simulating data $(X_i^{(k)}, Y_i^{(k)})$, $i = 1, \dots, n$ from the bivariate exponential distribution with $\lambda_1 = 1$ and $\lambda_1 < \lambda_2$ and the combinations of δ and n , and repeating the above steps. The variance of U_L is computed in three ways; (i) exact null variance using (7), (ii) sample variance, and (iii) jackknife variance estimate (Appendix E). Further, the empirical size and power are computed and compared using the standardised test obtained using the three variances. Such comparison is possible for the test U_L and we use the findings to validate the use of the jackknife estimate of the variance.

Table 1 presents the results for the empirical size of the test. The sample mean of the size is close to 1 and the three variances, exact null, sample variance and jackknife variance estimate, are close in all cases and hence, the jackknife variance estimate can be used for further analysis. Table 2 shows the empirical power of the test. The empirical power obtained by using the exact null variance and the jackknife method are close. For a fixed δ the empirical power increases with increase in the sample size and as λ_2 goes away from λ_1 . It decreases as δ increases from 0 to 1 for a fixed sample size. The power varies from 80% to 64% for $\lambda_2 = 1.2$ and sample size 500 as δ increases from 0 to 1. Similar power is achieved for the lower sample sizes of 50 and 100 but λ_2 being 1.5 and 2, that is further away from λ_1 .

Table 1 Empirical size for the likelihood-ratio ordering test in the absence of censoring under Gumbel's bivariate exponential model (19): Test is based on the U-statistics $U_L(5)$

n	$\delta = 0$				$\delta = 0.2$				$\delta = 0.5$				$\delta = 1$											
	Mean	Var	Null var	Size	J size	Mean	Var	Null var	Size	J size	Mean	Var	Null var	Size	J size	Mean	Var	Null var	Size	J size				
25	1.000	0.026	0.027	0.028	0.050	0.058	0.997	0.033	0.030	0.032	0.059	0.062	0.999	0.036	0.035	0.036	0.044	0.059	1.007	0.041	0.041	0.042	0.058	0.066
50	0.998	0.014	0.013	0.014	0.056	0.058	0.997	0.016	0.015	0.015	0.048	0.052	0.999	0.018	0.017	0.018	0.052	0.054	0.994	0.021	0.020	0.021	0.049	0.054
100	0.997	0.007	0.007	0.007	0.044	0.046	1.002	0.008	0.008	0.008	0.056	0.056	1.001	0.009	0.009	0.009	0.049	0.052	0.998	0.011	0.01	0.01	0.048	0.051
200	0.998	0.003	0.003	0.003	0.046	0.046	0.999	0.004	0.004	0.004	0.054	0.054	0.999	0.005	0.004	0.004	0.056	0.056	0.997	0.005	0.005	0.005	0.046	0.046
500	1.000	0.001	0.001	0.001	0.051	0.052	1.000	0.002	0.002	0.002	0.053	0.052	0.998	0.002	0.002	0.002	0.05	0.051	0.998	0.002	0.002	0.002	0.040	0.040

$\delta = 0$ corresponds to independence of X and Y . Both λ_1 , and λ_2 are one. The number of repeated samples was 2000. Mean and Var are sample mean and variance, Null var is the null variance obtained analytically, J var is jackknife variance estimate, J size is the size obtained by using jackknife variance and one sided test based on confidence limit (Appendix E). Size column gives the size of the test computed using the Null variance. The true size of the test was 0.05

Table 2 Empirical power for the likelihood-ratio ordering test in the absence of censoring under Gumbel's bivariate exponential model (19): Test is based on the U-statistics $U_L(5)$

λ_2	n	$\delta = 0$						$\delta = 0.2$						$\delta = 0.5$						$\delta = 1$					
		Mean	Var	Null var	J var	Power	J power	Mean	Var	Null var	J var	Power	J power	Mean	Var	Null var	J var	Power	J power	Mean	Var	Null var	J var	Power	J power
1.2	25	1.09	0.027	0.027	0.027	0.136	0.142	1.093	0.032	0.030	0.031	0.139	0.148	1.094	0.035	0.035	0.036	0.123	0.136	1.088	0.041	0.041	0.041	0.116	0.126
1.2	50	1.093	0.013	0.013	0.013	0.208	0.208	1.089	0.014	0.015	0.015	0.179	0.184	1.093	0.017	0.017	0.017	0.177	0.184	1.093	0.020	0.020	0.020	0.154	0.162
1.2	100	1.093	0.007	0.007	0.007	0.298	0.304	1.089	0.008	0.008	0.008	0.262	0.268	1.097	0.009	0.009	0.009	0.257	0.255	1.092	0.010	0.010	0.010	0.228	0.234
1.2	200	1.091	0.003	0.003	0.003	0.460	0.469	1.093	0.004	0.004	0.004	0.432	0.435	1.09	0.004	0.004	0.004	0.390	0.392	1.091	0.005	0.005	0.005	0.358	0.365
1.2	500	1.090	0.001	0.001	0.001	0.803	0.803	1.090	0.001	0.002	0.002	0.751	0.753	1.092	0.002	0.002	0.002	0.702	0.704	1.090	0.002	0.002	0.002	0.638	0.639
1.5	25	1.196	0.026	0.027	0.027	0.334	0.341	1.200	0.032	0.030	0.030	0.320	0.333	1.195	0.035	0.035	0.034	0.273	0.297	1.200	0.037	0.041	0.040	0.248	0.262
1.5	50	1.201	0.012	0.013	0.013	0.545	0.542	1.197	0.014	0.015	0.015	0.490	0.494	1.199	0.017	0.017	0.017	0.454	0.462	1.192	0.020	0.020	0.020	0.382	0.395
1.5	100	1.199	0.006	0.007	0.006	0.796	0.801	1.200	0.007	0.008	0.007	0.745	0.744	1.199	0.008	0.009	0.008	0.702	0.710	1.198	0.009	0.010	0.010	0.649	0.655
1.5	200	1.199	0.003	0.003	0.003	0.962	0.962	1.200	0.004	0.004	0.004	0.949	0.949	1.199	0.004	0.004	0.004	0.917	0.919	1.201	0.005	0.005	0.005	0.884	0.886
1.5	500	1.200	0.001	0.001	0.001	1.000	1.000	1.200	0.001	0.002	0.001	1.000	1.000	1.200	0.002	0.002	0.002	1.000	1.000	1.200	0.002	0.002	0.002	0.999	0.999
2.0	25	1.331	0.024	0.027	0.024	0.657	0.667	1.333	0.026	0.030	0.027	0.630	0.633	1.337	0.032	0.035	0.031	0.568	0.595	1.336	0.035	0.041	0.036	0.517	0.541
2.0	50	1.332	0.012	0.013	0.012	0.900	0.898	1.330	0.013	0.015	0.013	0.879	0.876	1.335	0.015	0.017	0.015	0.840	0.843	1.332	0.018	0.020	0.018	0.773	0.782
2.0	100	1.336	0.006	0.007	0.006	0.992	0.993	1.338	0.006	0.008	0.007	0.993	0.992	1.332	0.007	0.009	0.008	0.981	0.981	1.332	0.009	0.010	0.009	0.954	0.955
2.0	200	1.335	0.003	0.003	0.003	1.000	1.000	1.331	0.003	0.004	0.003	1.000	1.000	1.333	0.004	0.004	0.004	1.000	1.000	1.335	0.004	0.005	0.004	0.999	1.000
2.0	500	1.334	0.001	0.001	0.001	1.000	1.000	1.334	0.001	0.002	0.001	1.000	1.000	1.335	0.002	0.002	0.001	1.000	1.000	1.333	0.002	0.002	0.002	1.000	1.000

$\delta = 0$ corresponds to independence of X and Y . $\lambda_1 = 1$ and the number of repeated samples was 2000. Mean and Var are sample mean and variance, Null var is the null variance obtained analytically, J var is jackknife variance estimate. J power is the power obtained by using jackknife variance based test (Appendix E). Power column gives the power of the test computed using the Null variance. The size of the test was 0.05

Table 3 Empirical size and power for the likelihood-ratio ordering test in the presence of 20% censoring under Gumbel’s bivariate exponential model (19); Jackknife variance based tests using U-statistics U_{0c} , \widehat{U}_{2c} , \widehat{U}_{1c} , Eqs. (12), (15), (18) are applied (Appendix E)

λ_2	n	U-test	$\delta = 0$	$\delta = 0.5$	$\delta = 1$
1.0	100	U_{0c}	0.050	0.054	0.059
1.0	100	\widehat{U}_{2c}	0.035	0.039	0.043
1.0	100	\widehat{U}_{1c}	0.043	0.048	0.039
1.0	200	U_{0c}	0.046	0.049	0.053
1.0	200	\widehat{U}_{2c}	0.035	0.038	0.039
1.0	200	\widehat{U}_{1c}	0.049	0.045	0.037
1.5	100	U_{0c}	0.742	0.637	0.592
1.5	100	\widehat{U}_{2c}	0.566	0.473	0.417
1.5	100	\widehat{U}_{1c}	0.638	0.574	0.544
1.5	200	U_{0c}	0.940	0.885	0.851
1.5	200	\widehat{U}_{2c}	0.848	0.760	0.703
1.5	200	\widehat{U}_{1c}	0.908	0.832	0.804
2.0	100	U_{0c}	0.989	0.962	0.945
2.0	100	\widehat{U}_{2c}	0.920	0.860	0.815
2.0	100	\widehat{U}_{1c}	0.967	0.946	0.912
2.0	200	U_{0c}	1.000	1.000	0.999
2.0	200	\widehat{U}_{2c}	0.999	0.993	0.984
2.0	200	\widehat{U}_{1c}	1.000	0.998	0.998

$\delta = 0$ corresponds to independence of X and Y . Rows corresponding to $\lambda_2 = 1$ gives the empirical size while other values of λ_2 give the power. The number of repeated samples was 2000 and the size of the tests were 0.05

4.2 Censoring: performance of U_{0c} , \widehat{U}_{1c} and \widehat{U}_{2c}

In order to evaluate and compare the performance of U_{0c} , \widehat{U}_{1c} and \widehat{U}_{2c} , 20% censoring is considered. The censoring variable is assumed to be exponentially distributed with hazard λ_c , determined in order to achieve the desired proportion of censoring. The variances of the tests depend on the bivariate distribution as well as the censoring distribution, and are complicated to compute. Hence, we use the jackknife method to obtain empirical size and power (Appendix E). We also implemented bootstrap method to estimate the variance and carry out the hypothesis test based on confidence intervals. The number of bootstrap samples required for this are in the order of thousands (Efron and Stein 1981). Results from the two methods were comparable, however, the bootstrap method was too time consuming for large sample sizes compared to the jackknife method. Hence, in the sequel we present results using only the jackknife method.

Table 3 presents the simulation results of the empirical size and power of U_{0c} , \widehat{U}_{2c} and \widehat{U}_{1c} . The empirical sizes of tests U_{0c} and \widehat{U}_{1c} are close to the nominal level, whereas \widehat{U}_{2c} is slightly conservative. The power increases with increase in the sample size and decreases with increase in δ . The test U_{0c} performs the best both in terms of achieving the desired size and power, followed by \widehat{U}_{1c} and then \widehat{U}_{2c} .

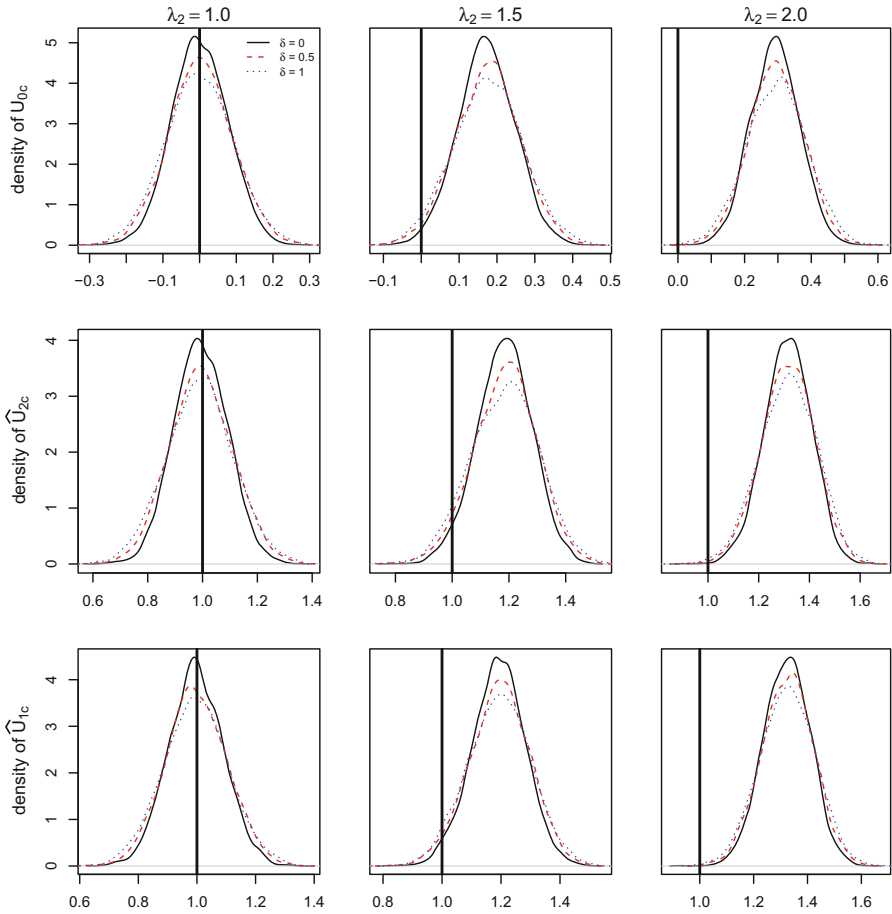


Fig. 1 Empirical densities of the U-statistics $U_{0c}, \hat{U}_{2c}, \hat{U}_{1c}$ for sample size 100 and 20% censoring. The number of repeated samples was 5000. The columns represent $\lambda_2 = 1, 1.5,$ and $2,$ and the rows are the three U-statistics, $U_{0c}, \hat{U}_{2c}, \hat{U}_{1c}$. The vertical line in each plot corresponds to the expectation of the test statistic under the null hypothesis

We further investigate influence of the proportion of censoring on the three tests by computing their empirical densities for $n = 100, \delta = (0, 0.5, 1)$ and three censoring proportions (20%, 50%, 80%). The samples are generated 5000 times for each combination of parameters. Figures 1, 2 and 3 exhibit the empirical densities for the nine combinations of varying λ_2 and δ . The empirical densities are bell shaped and center around the expected null value for $\lambda_2 = 1$ (left column in Fig. 1), and centered around larger values for $\lambda_2 = 1.5$ and 2 for sample size 100 and 20% censoring. The density of the test U_{0c} remains bell shaped and centered around the expected null value for 50% and 80% censoring. The density of the test \hat{U}_{1c} and \hat{U}_{2c} appear bell shaped and centered around the value slightly lower than the expected null value for 80% censoring. The test \hat{U}_{2c} performs the worst among the three tests. The overall effect of δ is not as dramatic as that of censoring. In most plots, the three empirical density curves

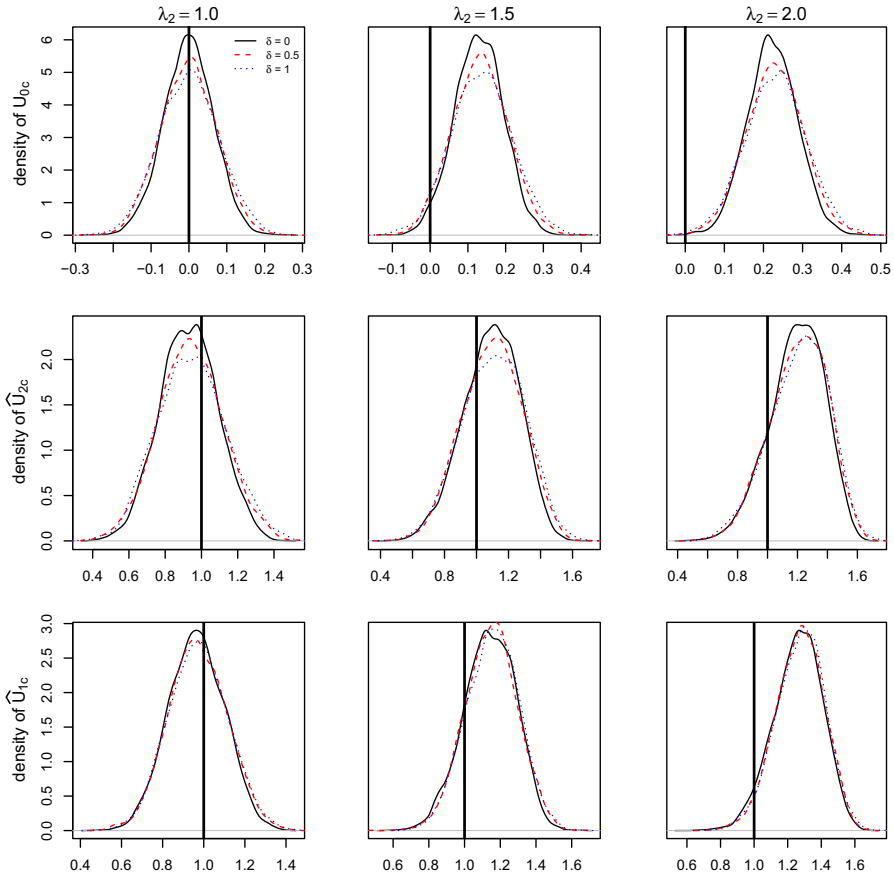


Fig. 2 Empirical densities of the U-statistics $U_{0c}, \widehat{U}_{2c}, \widehat{U}_{1c}$ for sample size 100 and 50% censoring. The number of repeated samples was 5000. The columns correspond to $\lambda_2 = 1, 1.5,$ and 2 and the rows correspond to the three U-statistics, $U_{0c}, \widehat{U}_{2c}, \widehat{U}_{1c}$. The vertical line in each plot corresponds to the expectation of the test statistic under the null hypothesis

overlap but the curve corresponding to $\delta = 0$ has the highest peak followed by 0.5 and then by 1. Similar observations can be made based on the three curves with regard to the shift away from the expected null for $\lambda_2 = 1.5$ and 2 . We further explored the effect of sample size and 50% censoring. Figures 4 and 5 show the density plots for $n = 500$ and 1000 , respectively. It is apparent that the empirical densities of \widehat{U}_{2c} and \widehat{U}_{1c} appear bell shaped and center to the right of the expected null for sample size 500 (Fig. 4) but larger shift for $n = 1000$ (Fig. 5).

We also compute the statistics U_{1c} and U_{2c} for known $S_c(\cdot)$ to examine the effect of weights on the spread of the distribution. Figures 6 and 7 present densities of U_{1c} and U_{2c} for the censoring proportion 50%, and sample sizes 500 and 1000. The exponential censoring distribution with hazard λ_c was used to compute the U-statistics with known censoring distribution (solid lines) and the Kaplan-Meier estimate is used as before (dashed lines). It is apparent from all plots that the tails of the solid curves

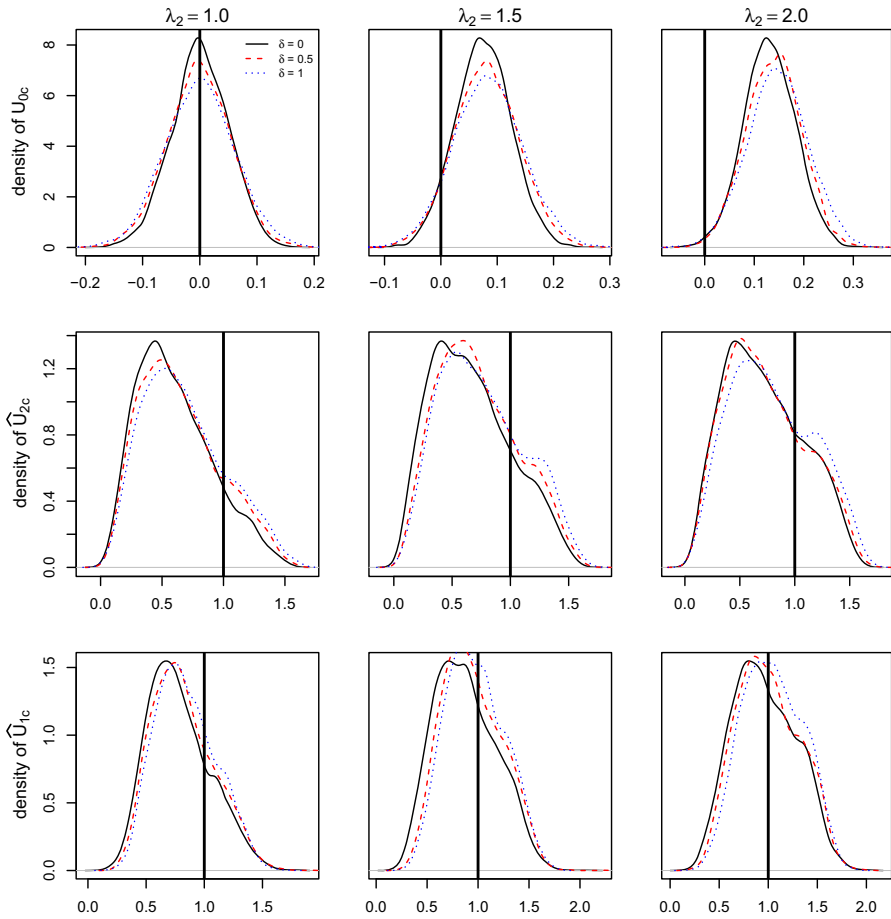


Fig. 3 Empirical densities of the U-statistics U_{0c} , \widehat{U}_{2c} , \widehat{U}_{1c} for sample size 100 and 80% censoring. The number of repeated samples was 5000. The columns correspond to $\lambda_2 = 1, 1.5,$ and 2 and the rows correspond to the three U-statistics, U_{0c} , \widehat{U}_{2c} , \widehat{U}_{1c} . The vertical line in each plot corresponds to the expectation of the test statistic under the null hypothesis

are long compared to the dashed ones. Also, the spread of the densities with dashed line is narrower than the ones with solid line. This particularly explains relatively poor performance of \widehat{U}_{1c} and \widehat{U}_{2c} .

5 Applications

The proposed tests are applied to three real data sets. The first data has small sample size and helps to visualise the amount of weights given to the observations in the case of censoring. The second data has heavy censoring and is used for illustrating the differences in the findings of the proposed tests. The third data are complete and has large sample size, allowing experimenting with different proportions of censoring.

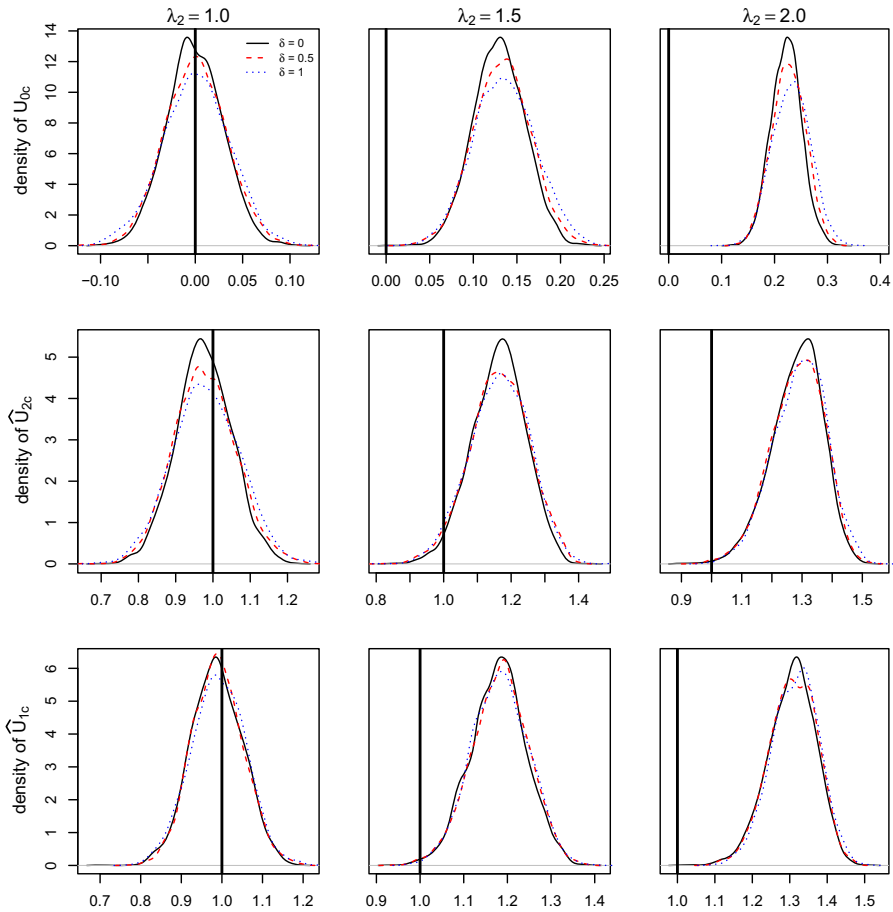


Fig. 4 Empirical densities of the U-statistics U_{0c} , \widehat{U}_{2c} , \widehat{U}_{1c} for sample size 500 and 50% censoring. The number of repeated samples was 5000. The columns correspond to $\lambda_2 = 1, 1.5,$ and 2 and the rows correspond to the three U-statistics, U_{0c} , \widehat{U}_{2c} , \widehat{U}_{1c} . The vertical line in each plot corresponds to the expectation of the test statistic under the null hypothesis

5.1 Skin graft data

The skin graft data, though smaller in size, has been widely analysed (Lin and Ying 1993) and has only 18% of censoring. The data consist of survival times (X, Y) , in days, of closely and poorly matched skin grafts on the same burn patient. Both (X, Y) are subjected to the independent univariate censoring. The data include only 11 patients and the survival times of two closely matched grafts are censored (18%). There is no censoring in the Y observations. Here U_{0c} and its jackknife variance estimate are 0.409 and 0.033. The 90% confidence interval for the expectation of U_{0c} is (0.109, 0.709), and the lower confidence limit is above the null expectation of 0. The Kaplan-Meier estimate of the censoring survival function has jumps only at 57 and 63 days, and the values of jumps are 0.75 and 0.50, respectively. The \widehat{U}_{2c} test value is 1.23 and the

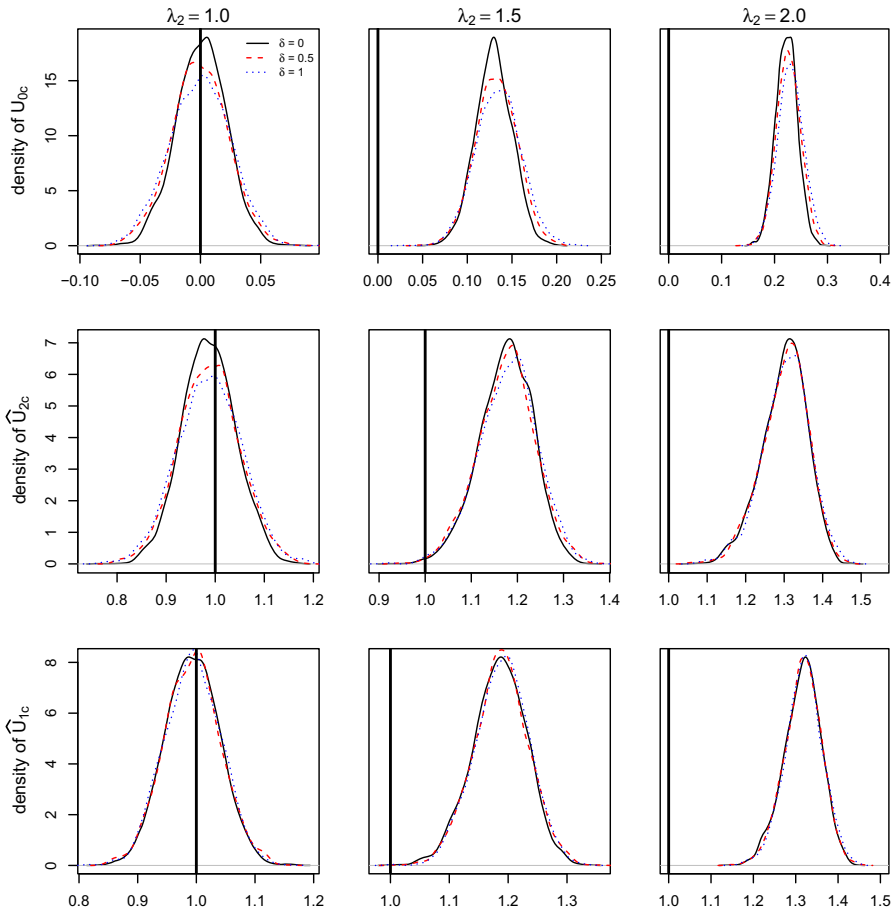


Fig. 5 Empirical densities of the U-statistics $U_{0c}, \hat{U}_{2c}, \hat{U}_{1c}$ for sample size 1000 and 50% censoring. The number of repeated samples was 5000. The columns correspond to $\lambda_2 = 1, 1.5,$ and 2 and the rows correspond to the three U-statistics, U_{0c}, U_{2c}, U_{1c} . The vertical line in each plot corresponds to the expectation of the test statistic under the null hypothesis

confidence interval is (0.92, 1.55). The \hat{U}_{1c} test provides similar result, the value of the test statistic is 1.40 and the confidence interval is (1.10, 1.69). The expected null value 1 falls below the lower confidence limit in case of \hat{U}_{1c} but not of \hat{U}_{2c} . Note that tests U_{0c} and \hat{U}_{1c} utilises more information compared to \hat{U}_{2c} . This is a good data to illustrate how the three tests for censored case compare since the proportion of censoring is low, even though the sample size is small. The Kaplan-Meier survival curve for X dominates that of Y (Fig. 8) implying that X is stochastically greater than Y . This is a necessary condition for X being greater than Y in the sense of likelihood ratio (Shanthikumar et al. 1991).

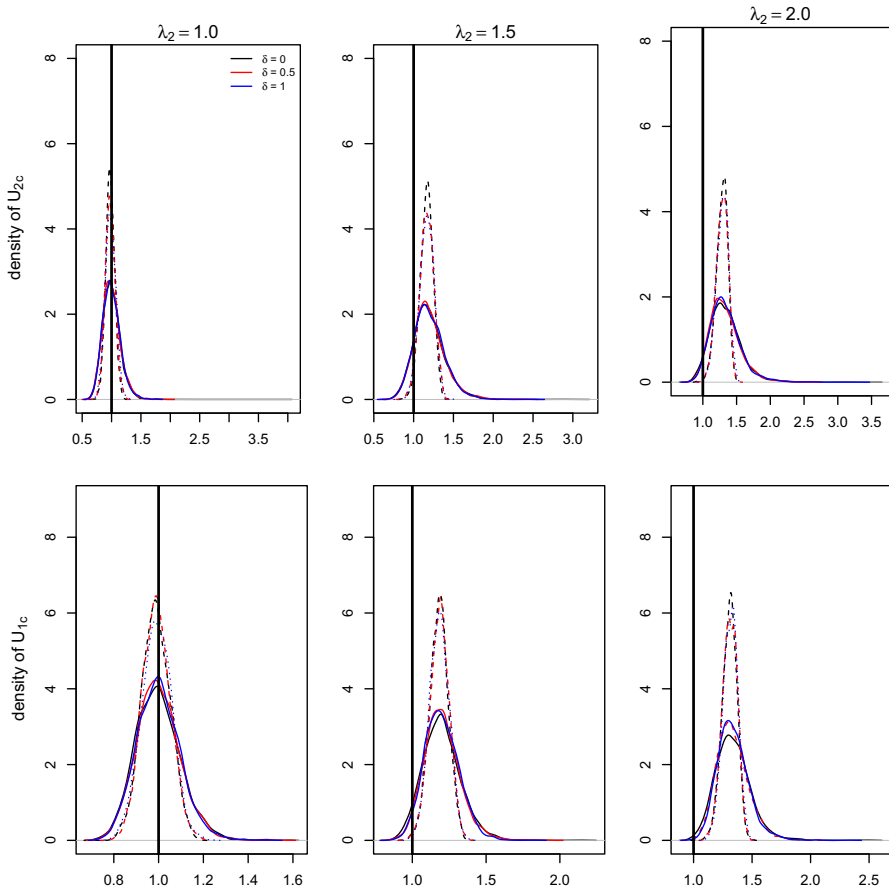


Fig. 6 Empirical densities of the U-statistics (U_{2c} , \widehat{U}_{2c}) and (U_{1c} , \widehat{U}_{1c}) for sample size 500 and 50% censoring. Solid and dashed lines represent densities of an U-statistic evaluated with known and estimated censoring distribution, respectively. The columns correspond to $\lambda_2 = 1, 1.5,$ and 2 and the rows correspond to the two U-statistics, U_{2c} and U_{1c} . The vertical line in each plot corresponds to the expectation of the test statistic under the null hypothesis

5.2 Diabetic retinopathy data

The diabetic retinopathy data consist of follow-up of 197 diabetic patients for a fixed period of time. Each patient has one eye randomized to the laser treatment and the other eye receives no treatment. For each eye, the event of interest is the time from initiation of treatment to the time when visual acuity drops below 5/200 for two visits in a row. For both eyes, times to the visual acuity drop, in months, are recorded. In the present analysis, X and Y are the times to visual acuity drop for the treated and for the untreated eye, respectively. There is 73% and 49% censoring in the treated and untreated eyes, respectively. Only 19% data are observed in both eyes, i.e., proportion of censoring is 81% and 41% is censored in both. The data are available from <https://www.mayo.edu/research/documents/diabeteshtml/DOC-10027460>.

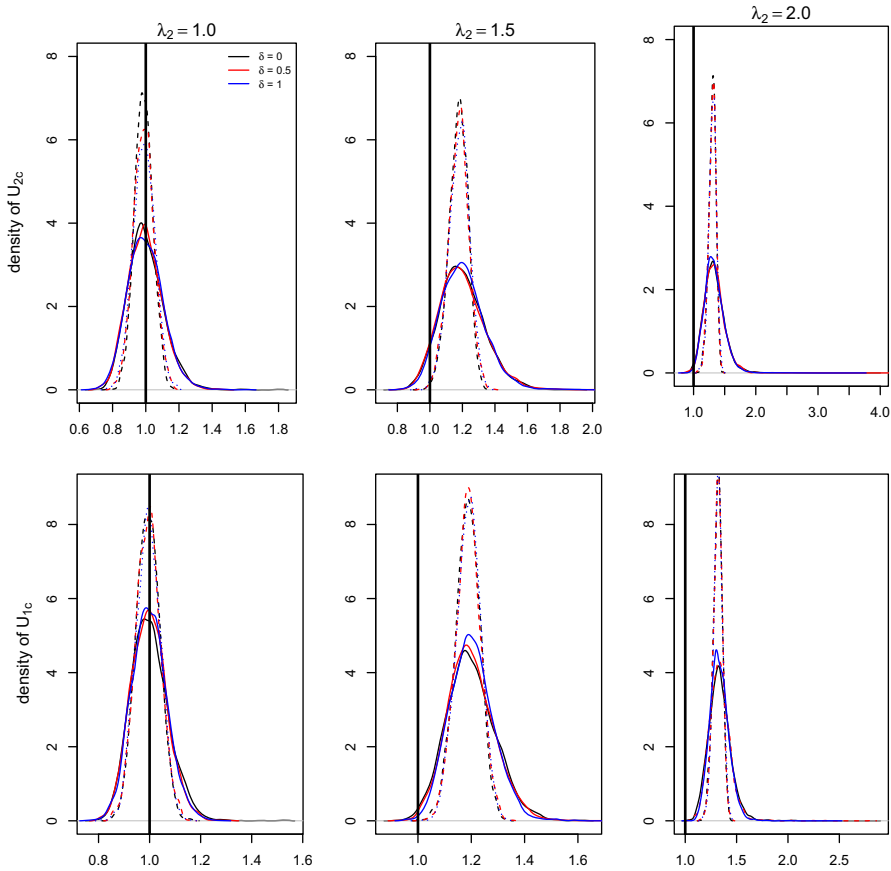


Fig. 7 Empirical densities of the U-statistics (U_{2c}, \widehat{U}_{2c}) and (U_{1c}, \widehat{U}_{1c}) for sample size 1000 and 50% censoring. Solid and dashed lines represent densities of an U-statistic evaluated with known and estimated censoring distribution, respectively. The columns correspond to $\lambda_2 = 1, 1.5,$ and 2 and the rows correspond to the two U-statistics, U_{2c}, U_{1c} . The vertical line in each plot corresponds to the expectation of the test statistic under the null hypothesis

In this example, the Kaplan-Meier estimate of the censoring survival function is based on 159 censoring observations, and the last observation is *uncensored*. The value and the 90% confidence interval of the U_{0c} test are 0.20 and (0.14, 0.27). For \widehat{U}_{2c} and \widehat{U}_{1c} , these are 0.06, (0.03, 0.10) and 0.19, (0.12, 0.27), respectively. The U_{0c} test suggests that the diabetes data are extreme for the null hypothesis of equality of the two marginal densities against the LR ordering $f(t)/g(t)$ is nondecreasing, while the other two tests suggests that $f(t)/g(t)$ is nonincreasing. The Kaplan-Meier curves for the marginal survival functions of X and Y crossed and hence, there may not be stochastic ordering between X and Y . This in turn means that there may not be likelihood-ratio ordering since the necessary condition may not be true.

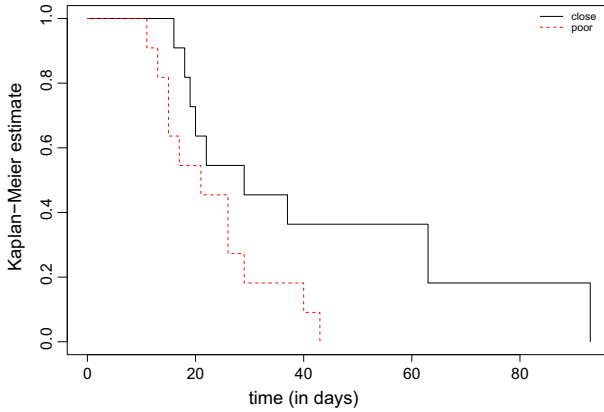


Fig. 8 Kaplan-Meier estimates of survival times of closely (X) and poorly (Y) matched skin grafts

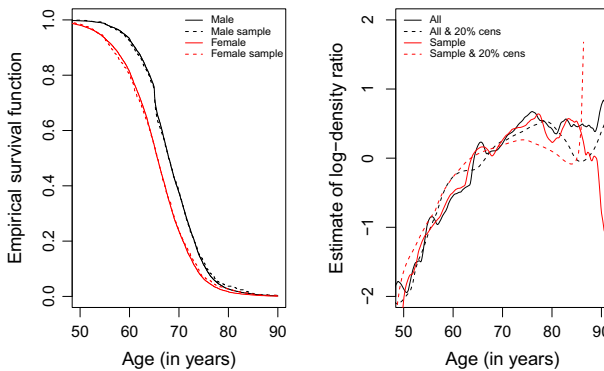


Fig. 9 Empirical estimates of ages of couples at the time of entry to the study, $X =$ age of a male and $Y =$ age of his female partner. The left panel shows the empirical survival functions and the right panel shows the estimates of the log-density ratio. The estimates are based on the original data of size 14884 and a sample of size 1000 selected from it without and with 20% independent univariate censoring imposed

5.3 Canadian insurance data set

The insurance data on couples have been analysed by several researchers and we refer to Frees et al. (1996), Luciano et al. (2008) for the detailed description of the data. Here the ages of couples, at least 40 years of age, at the start of the follow-up are taken as X (male) and Y (female). The data consist of 14884 couples. Figure 9 shows the marginal empirical estimates of the survival functions of X and Y (solid line, left panel) and it is clear that the survival function of the age of the male dominates that of the age of the female partner. An estimate of the kernel type estimate of the log-ratio of the marginal densities (black solid line, right panel, Fig. 9) exhibits a nondecreasing trend. The value and confidence intervals of U_L test for uncensored data are 1.248 and (1.242, 1.254) suggesting that the couples data may be an outlier for testing the hypothesis of equal densities against LR ordering.

The censoring times are simulated from an uniform distribution over the interval (40, b). The parameter b is determined so that the desired proportion of censoring, $(1 - p)$ can be achieved. It is easy to see that $b = (\mu - 40 * p)/(1 - p)$ where μ is the expectation of X or Y . First, the censoring times are simulated to achieve 20% censoring. The kernel type density estimates for the marginal densities are plotted in Fig. 9 (black dashed line, right panel). A nondecreasing trend can be observed in this case. With 20% and 50% censoring, the value of the test U_{0c} and the confidence interval for its expectation are 0.180, (0.175, 0.186), and 0.115, and (0.112, 0.120), respectively. Notice that the lower confidence limit is above 0 and hence, the conclusion is the same as in the uncensored case. The values of \widehat{U}_{1c} , (1.2470, 1.2479), and \widehat{U}_{2c} , (1.248, 1.246), are similar for 20% and 50% censoring and are larger than 1. Due to the large sample size, the jackknife method is time consuming and hence, is not implemented to obtain the confidence interval.

A sample of size 1000 is drawn from the original data to compare different tests. Figure 9 also shows the marginal empirical estimates of the survival functions (dashed line, left panel) and the estimate of the log-density ratio of X and Y for the sample (red solid line, right panel). The two dashed curves are close to the solid curves indicating that the sample is a representative of the original data. The estimate of log-density ratio based on the sample also shows nondecreasing trend. The value and confidence intervals of $E(U_L)$ for the sample are 1.238, and (1.216, 1.260), which are close to the results obtained for the original data. This sample is also subject to 20% independent univariate censoring and analysed using tests for censored data as above. The estimate of the log-density ratio is computed as before and displayed in Fig. 9 (red dashed line, right panel). In this case also a general nondecreasing trend can be seen. With 20% censoring imposed on the sample, the values of the U_{0c} , \widehat{U}_{1c} and \widehat{U}_{2c} and the confidence intervals are: 0.176, (0.158, 0.195), 1.238, (1.209, 1.267) and 1.232, (1.206, 1.258). For 50% censoring, these are 0.104, (0.091, 0.118), 1.223, (1.182, 1.263) and 1.219, (1.187, 1.252). In this case, all tests provide consistent results indicative of LR ordering between the ages of insured couples.

6 Discussion

Tests based on U-statistics for testing equality of the marginal densities against likelihood-ratio ordering for bivariate data have been proposed in this paper. The test for complete data maintains its size and has good power for a sample of size n at least 50, and a large choice of parameters of Gumbel's bivariate exponential family of distributions. The value of the dependence parameter δ has no major role to play. The tests for censored data merit more discussion. The simulation results show that the performance of the tests depend on n , δ and the percentage of censoring. For sample size in the range of 1000 all tests perform well irrespective of the percentage of censoring and the dependency. One may need larger sample size to achieve the same power for highly dependent pairs compared to independent pairs.

Of the tests proposed for the censoring scenario, the test based on U_{0c} performs the best both in terms of achieving the desired size and power, followed by \widehat{U}_{1c} and then \widehat{U}_{2c} . Both \widehat{U}_{2c} and \widehat{U}_{1c} are weighted U-statistics, and they use the Kaplan-Meier

estimate of the censoring distribution as weights, where higher weights are given to late failures compared to the true weights. Moreover, the Kaplan-Meier estimate remains constant after the last observed censoring time. On the other hand U_{0c} is a U-statistic which is simple, easy to compute, and makes a better use of the available data. The U-statistic has limiting normal distribution for relatively small values of n .

It is interesting to note that the expectation of U_{0c} depends on the censoring survival function $S_c()$, whereas the choice of weight functions ensures that the expectations of U_{1c} and U_{2c} are independent of $S_c()$. However, the expectations of \widehat{U}_{1c} and \widehat{U}_{2c} depend on unknown $S_c()$. It is known that the Kaplan-Meier estimator is biased upward (page 257, Andersen et al. 1993) but asymptotically unbiased. Similarly, the expectations of \widehat{U}_{1c} and \widehat{U}_{2c} are biased. Moreover, the bias and the variance would depend on the sample size, the percentage of censoring and the value of dependency parameter. These need detailed investigation.

We have carried out simulation studies for Gumbel's bivariate exponential family which models a negative association between the two variables. In addition, simulation studies have been performed for insurance data, where the variables of interest have positive association (Kendall's τ being 0.56). A graphical check for monotonicity of log-density ratio based on kernel type estimates of marginal densities has been illustrated for insurance data using complete as well as censored data. This gives a crude method to visually check monotonicity of ratio of densities.

The three tests gave inconsistent results for the diabetes data. Due to heavy censoring, the findings of the tests may not be valid and modifications to the tests for heavy censoring are needed. Possible extensions of the proposed tests to allow for variation in sampling and censoring schemes are needed for their wider applicability. The sampling scheme which results into left truncation and late entry would generate truncated data, where the observed data may not be even identically distributed. Other types of censoring that arise in practice and hence, are of interest: univariate dependent censoring, bivariate censoring with independent censoring times which could be drawn from the same distribution/ different distributions/bivariate distribution. There may be natural ordering between X and Y so that $X < Y$. For example, in the Stanford study of post-heart transplant deaths, the transplant time precedes the death time (Kalbfleisch and Prentice 2002). Moreover, both times are censored if there is death before the transplant. In this case the censoring may not be independent. Similar examples arise in social sciences, e.g., ages of a female at the time of stopping formal education and marriage are usually ordered.

Acknowledgements We wish to thank the Society of Actuaries, through the courtesy of Edward (Jed) Frees and Emiliano A. Valdez, for providing the Canadian data used in this paper. We thank the reviewers for their critical comments.

Funding Open Access funding provided by University of Helsinki including Helsinki University Central Hospital.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If

material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

A: Asymptotic distribution of U_{0c} (12)

In order to obtain the asymptotic variance of the U-statistic (12), we first note that for fixed $z_i = (x_i, y_i)$,

$$\begin{aligned}\phi_{0c1}(z_i) &= E(\phi_{0c}(Z_i, Z_j)|z_i) \\ &= \frac{1}{2} \left[P(\delta_j^y = 1, U_j < t_i) + I(\delta_i^y = 1)P(u_i < T_j) \right. \\ &\quad \left. - I(\delta_i^x = 1)P(U_j > t_i) - P(\delta_j^x = 1, u_i > T_j) \right].\end{aligned}$$

The variance of $\phi_{0c1}(Z_i)$ is $\sigma_{0c1}^2 = E(\phi_{0c1}^2(Z_i)) - (E(\phi_{0c1}(Z_i)))^2$, where the expectations are

$$E(\phi_{0c1}(Z_i)) = \left[\int_0^\infty S_c^2(y) \bar{F}(y) g(y) dy - \int_0^\infty S_c^2(x) \bar{G}(x) f(x) dx \right],$$

as in (13), and $E(\phi_{0c1}^2(Z_i))$ can be obtained similarly.

The following theorem follows from Hoeffding's decomposition (Hoeffding 1948; Lee 2020).

Theorem 2 *If $E(\phi_{0c}^2(\tilde{Z}_i, \tilde{Z}_j)) < \infty$ and $\sigma_{0c1}^2 > 0$ then $\sqrt{n}(U_{0c} - E(U_{0c}))$ converges in distribution to $N(0, 4\sigma_{0c1}^2)$ as $n \rightarrow \infty$. Under the null hypothesis $E(U_{0c}) = 0$.*

B: Asymptotic distribution of U_{2c} (14) and \hat{U}_{2c} (15)

We assume that the censoring distribution is known and obtain the asymptotic distribution of U_{2c} . Let $\phi_{2c}(\tilde{Z}_i, \tilde{Z}_j)$ be the kernel used to define U_{2c} in (14). Let $\phi_{2c1}(\tilde{z}_i)$ be its expectation obtained by conditioning on \tilde{Z}_i .

$$\begin{aligned}\phi_{2c1}(\tilde{z}_i) &= E \left((I(U_j < t_i) + I(u_i < T_j)) \frac{\delta_i^x \delta_i^y \delta_j^x \delta_j^y}{S_c(t_i \vee u_i -) S_c(T_j \vee U_j -)} \right) \\ &= \frac{\delta_i^x \delta_i^y}{S_c(t_i \vee u_i -)} E \left((I(U_j < t_i) + I(u_i < T_j)) \frac{\delta_j^x \delta_j^y}{S_c(T_j \vee U_j -)} \right) \\ &= \frac{\delta_i^x \delta_i^y}{S_c(t_i \vee u_i -)} \left[G(t_i) + \bar{F}(u_i) \right].\end{aligned}\tag{20}$$

The above expectation is with respect to the joint distribution of $(T_j, \delta_j^x = 1, U_j, \delta_j^y = 1)$. This joint distribution is given by

$$\begin{aligned}
 P(T \leq x, \delta^x = 1, U \leq y, \delta^y = 1) &= P(X \leq x, C > X, Y \leq y, C > Y) \\
 &= \int_0^x \int_0^y S_c(t \vee u)w(t, u)dudt,
 \end{aligned}
 \tag{21}$$

and the joint density function is $S_c(x \vee y)w(x, y)$. Note that $E(\phi_{2c1}(\tilde{Z}_i))$ w.r.t. the joint distribution of $(T_i, \delta_i^x = 1, U_i, \delta_i^y = 1)$ is the same as $E(U_{2c})$ which is $\Delta(F, G)$. The variance of ϕ_{2c1} is $\sigma_{2c1}^2 = E(\phi_{2c1}^2(\tilde{Z}_i)) - (E(\phi_{2c1}(\tilde{Z}_i)))^2$. Here the expectation is w.r.t. the joint density of $(T, \delta_x = 1, U, \delta_y = 1)$.

Theorem 3 *If $E(\phi_{2c}^2(\tilde{Z}_i, \tilde{Z}_j)) < \infty$ and $\sigma_{2c1}^2 > 0$ then $\sqrt{n}(U_{2c} - E(U_{2c}))$ converges in distribution to $N(0, 4\sigma_{2c1}^2)$ as $n \rightarrow \infty$. Under the null hypothesis $E(U_{2c}) = 1$ and σ_{2c1}^2 depends on (F, G) and the censoring distribution.*

Proof The asymptotic normality of the U-statistic follows from Hoeffding’s decomposition (Hoeffding 1948; Lee 2020). Note that the expectation of the square of $\phi_{2c1}^2(\tilde{Z}_i)$ is

$$\begin{aligned}
 E(\phi_{2c1}^2(\tilde{Z}_i)) &= \int_0^\infty \int_0^\infty \frac{G^2(x) + \bar{F}^2(y)}{S_c(x \vee y-)} w(x, y)dydx \\
 &\quad + 2 \int_0^\infty \int_0^\infty \frac{G(x)\bar{F}(y)}{S_c(x \vee y-)} w(x, y)dydx
 \end{aligned}$$

Under the null hypothesis $f(\cdot) = g(\cdot)$, (20) simplifies.

$$\begin{aligned}
 \phi_{2c1}(\tilde{z}_i) &= \frac{\delta_i^x \delta_i^y}{S_c(t_i \vee u_i-)} \left[F(t_i) + \bar{F}(u_i) \right], \\
 E(\phi_{2c1}(\tilde{Z}_i)) &= 1, \text{ and} \\
 E(\phi_{2c1}^2(\tilde{Z}_i)) &= \int_0^\infty \int_0^\infty \frac{F^2(x) + \bar{F}^2(y)}{S_c(x \vee y-)} w(x, y)dydx \\
 &\quad + 2 \int_0^\infty \int_0^\infty \frac{F(x)\bar{F}(y)}{S_c(x \vee y-)} w(x, y)dydx.
 \end{aligned}$$

The variance σ_{2c1}^2 clearly depends on (F, G) and the censoring distribution. □

We sketch the proof of asymptotic normality of $\sqrt{n}(\hat{U}_{2c} - E(U_{2c}))$ below. Recall that an estimator \hat{U}_{2c} of U_{2c} is obtained by replacing the survival function $S_c(t)$ of the censoring time by the Kaplan-Meier estimator $\hat{S}_c(t)$. We first note that

$$\begin{aligned} \sqrt{n}(\widehat{U}_{2c} - E(U_{2c})) &= \sqrt{n}(U_{2c} - E(U_{2c})) \\ &+ \sqrt{n} \frac{1}{\binom{n}{2}} \sum_{1 \leq i < j \leq n} \phi(Z_i, Z_j) \delta_i^x \delta_i^y \delta_j^x \delta_j^y \\ &\left[\frac{1}{\widehat{S}_c(T_i \vee U_{i-}) \widehat{S}_c(T_j \vee U_{j-})} - \frac{1}{S_c(T_i \vee U_{i-}) S_c(T_j \vee U_{j-})} \right]. \end{aligned} \tag{22}$$

The term in the square brackets can be simplified as:

$$\begin{aligned} &\left[\frac{1}{\widehat{S}_c(T_i \vee U_{i-}) \widehat{S}_c(T_j \vee U_{j-})} - \frac{1}{S_c(T_i \vee U_{i-}) S_c(T_j \vee U_{j-})} \right] \\ &= - \left[\frac{(\widehat{S}_c(T_i \vee U_{i-}) - S_c(T_i \vee U_{i-}))}{\widehat{S}_c(T_i \vee U_{i-}) S_c(T_i \vee U_{i-}) S_c(T_j \vee U_{j-})} + \frac{(\widehat{S}_c(T_j \vee U_{j-}) - S_c(T_j \vee U_{j-}))}{\widehat{S}_c(T_i \vee U_{i-}) \widehat{S}_c(T_j \vee U_{j-}) S_c(T_j \vee U_{j-})} \right]. \end{aligned} \tag{23}$$

Combining (22) and (23) we get an equation which resembles Eq. (24) in Datta et al. (2010). Further, each of the terms in the square brackets can be written as an integration with respect to the empirical subdistribution function $H_n^{uc}(t, u)$ ($T_i, \delta_i^x = 1, U_i, \delta_i^y = 1$);

$$H_n^{uc}(t, u) = \frac{1}{n} \sum_{i=1}^n I(T_i \leq t, \delta_i^x = 1, U_i \leq u, \delta_i^y = 1).$$

In general, for some kernel $K()$ the double integral w.r.t. the empirical distribution is

$$\begin{aligned} &\int \int K(t_1, u_1, t_2, u_2) \phi(t_1, u_1, t_2, u_2) dH_n^{uc}(t_1, u_1) dH_n^{uc}(t_2, u_2) \\ &= \frac{1}{n^2} \sum_{i=1}^n \sum_{j=1}^n K(T_i, U_i, T_j, U_j) \phi(T_i, U_i, T_j, U_j) \delta_i^x \delta_i^y \delta_j^x \delta_j^y \end{aligned}$$

Let $\lambda_c(u)$ be the hazard rate of the censoring variable C . Let $N_i^C(t) = I(L_i \leq t, \delta_i = 0)$ be the counting process corresponding to censoring for i th individual and $R_i(t) = I(L_i \geq t)$ be the at-risk process. The martingale associated with the counting process $N_i^C(t)$ with respect to the self-exciting history is given by

$$M_i^C(t) = N_i^C(t) - \int_0^t R_i(u) \lambda_c(u) du.$$

Let $N^C(t) = \sum_{i=1}^n N_i^C(t)$, $R(t) = \sum_{i=1}^n R_i(t)$ be the aggregated counting and at-risk process over n units. We can now use weak convergence of the Kaplan-Meier estimator, and the arguments given in Theorem 1 of Datta et al. (2010) to prove the asymptotic normality of \widehat{U}_{2c} . Note that the subdistribution function $W_n()$ in their theorem is replaced by the subdistribution function $H_n^{uc}()$.

C: Asymptotic distribution of U_{1c} (17) and \widehat{U}_{1c} (18)

We assume that the censoring distribution is known and obtain the asymptotic distribution of U_{1c} . Define $\phi_{1c1}(\tilde{z}_i)$, the expectation of $\phi_{1c}(\tilde{Z}_i, \tilde{Z}_j)$ by conditioning on \tilde{Z}_i .

$$\begin{aligned} \phi_{1c1}(\tilde{z}_i) &= E[\phi_{1c}(\tilde{Z}_i, \tilde{Z}_j) \mid \tilde{z}_i] \\ &= E\left(I(Y_j < x_i) \frac{\delta_i^x \delta_j^y}{S_c(x_i-)S_c(Y_j-)} + I(y_i < X_j) \frac{\delta_i^y \delta_j^x}{S_c(y_i-)S_c(X_j-)} \right) \\ &= \frac{I(x_i \leq c_i)}{S_c(x_i-)} E\left(\frac{I(Y_j < x_i)I(Y_j \leq C_j)}{S_c(Y_j-)} \right) \\ &\quad + \frac{I(y_i \leq c_i)}{S_c(y_i-)} E\left(\frac{I(X_j > y_i)I(X_j \leq C_j)}{S_c(X_j-)} \right) \\ &= \frac{I(x_i \leq c_i)}{S_c(x_i-)} \int_0^{x_i} g(y)dy + \frac{I(y_i \leq c_i)}{S_c(y_i-)} \int_{y_i}^\infty f(x)dx \\ &= \frac{I(x_i \leq c_i)}{S_c(x_i-)} G(x_i) + \frac{I(y_i \leq c_i)}{S_c(y_i-)} \bar{F}(y_i). \end{aligned} \tag{24}$$

Note that the expectation of $\phi_{1c1}(\tilde{Z}_i)$ w.r.t. the joint distribution of $(T_i, \delta_i^x = 1, U_i, \delta_i^y = 1)$ is the same as the expectation $E(U_{1c})$ which is $\Delta(F, G)$. The variance of $\phi_{1c1}(\tilde{Z}_i)$ is $\sigma_{1c1}^2 = E(\phi_{1c1}^2(\tilde{Z}_i)) - (E(\phi_{1c1}(\tilde{Z}_i)))^2$.

Theorem 4 *If $E(\phi_{1c}^2(\tilde{Z}_i, \tilde{Z}_j)) < \infty$ and $\sigma_{1c1}^2 > 0$ then $\sqrt{n}(U_{1c} - E(U_{1c}))$ converges in distribution to $N(0, 4\sigma_{1c1}^2)$ as $n \rightarrow \infty$. Under the null hypothesis $E(U_{1c}) = 1$ and σ_{1c1}^2 depends on (F, G) and the censoring distribution.*

Proof The asymptotic normality of the U-statistic follows from Hoeffding’s decomposition (Hoeffding 1948; Lee 2020). The variance σ_{1c1}^2 can be obtained by noting that

$$\begin{aligned} E(\phi_{1c1}^2(\tilde{Z}_i)) &= \int_0^\infty \frac{G^2(x)}{S_c(x-)} f(x)dx + \int_0^\infty \frac{\bar{F}^2(y)}{S_c(y-)} g(y)dy \\ &\quad + 2 \int_0^\infty \int_0^\infty \frac{G(x)\bar{F}(y)}{S_c(x-)S_c(y-)} S_c(x \vee y)w(x, y)dydx. \end{aligned}$$

Under the null hypothesis $f(\cdot) = g(\cdot)$, (24) simplifies and so does its expectation, and the expectation of its square.

$$\begin{aligned} \phi_{1c1}(\tilde{z}_i) &= \frac{I(x_i \leq c_i)}{S_c(x_i-)} F(x_i) + \frac{I(y_i \leq c_i)}{S_c(y_i-)} \bar{F}(y_i), \\ E(\phi_{1c1}(\tilde{Z}_i, \tilde{Z}_j)) &= E(\phi_{1c1}(\tilde{Z}_i)) = 1, \text{ and} \\ E(\phi_{1c1}^2(\tilde{Z}_i)) &= \int_0^\infty \left[\frac{F^2(x)}{S_c(x-)} + \frac{\bar{F}^2(x)}{S_c(x-)} \right] f(x)dx, \end{aligned}$$

$$+ 2 \int_0^\infty \int_0^\infty \frac{F(x)\bar{F}(y)}{S_c(x-)S_c(y-)} S_c(x \vee y)w(x, y)dydx.$$

It is obvious from the above expectations that σ_{1c1}^2 depends on (W, F, G) and the censoring distribution. □

Recall that an estimator \widehat{U}_{1c} of U_{1c} is obtained by replacing the survival function $S_c(t)$ of the censoring time by the Kaplan-Meier estimator $\widehat{S}_c(t)$. As in Appendix C, we first note that

$$\begin{aligned} \sqrt{n}(\widehat{U}_{1c} - E(U_{1c})) &= \sqrt{n}(U_{1c} - E(U_{1c})) \\ &+ \sqrt{n} \frac{1}{\binom{n}{2}} \sum_{1 \leq i < j \leq n} I(U_j < T_i) \delta_i^x \delta_j^y \left[\frac{1}{\widehat{S}_c(T_i-) \widehat{S}_c(U_j-)} - \frac{1}{S_c(T_i-) S_c(U_j-)} \right] \\ &+ \sqrt{n} \frac{1}{\binom{n}{2}} \sum_{1 \leq i < j \leq n} I(U_i < T_j) \delta_i^y \delta_j^x \left[\frac{1}{\widehat{S}_c(U_i-) \widehat{S}_c(T_j-)} - \frac{1}{S_c(U_i-) S_c(T_j-)} \right]. \end{aligned} \tag{25}$$

The second and third terms can be simplified by noting that:

$$\begin{aligned} &\left[\frac{1}{\widehat{S}_c(T_i-) \widehat{S}_c(U_j-)} - \frac{1}{S_c(T_i-) S_c(U_j-)} \right] \\ &= - \left[\frac{(\widehat{S}_c(T_i-) - S_c(T_i-))}{\widehat{S}_c(T_i-) S_c(T_i-) S_c(U_j-)} + \frac{(\widehat{S}_c(U_j-) - S_c(U_j-))}{\widehat{S}_c(T_i-) \widehat{S}_c(U_j-) S_c(U_j-)} \right]. \end{aligned} \tag{26}$$

Combining Eqs. (25) and (26) results into an expression similar to Eq. (24) in Datta et al. (2010). Further, each of the terms in the square brackets can be written as an integration with respect to the empirical subdistributions $H_{T_n}^{uc}(t)$ of $(T_i, \delta_i^x = 1)$ and $H_{U_n}^{uc}(u)$ of $(U_i, \delta_i^y = 1)$;

$$\begin{aligned} H_{T_n}^{uc}(t) &= \frac{1}{n} \sum_{i=1}^n I(T_i \leq t, \delta_i^x = 1), \\ H_{U_n}^{uc}(u) &= \frac{1}{n} \sum_{i=1}^n I(U_i \leq u, \delta_i^y = 1). \end{aligned}$$

In general, for some kernel $K()$ the double integral w.r.t. the empirical subdistributions is

$$\int \int K(t, u) \phi(t, u) dH_{T_n}^{uc}(t) dH_{U_n}^{uc}(u) = \frac{1}{n^2} \sum_{i=1}^n \sum_{j=1}^n K(T_i, U_j) \phi(T_i, U_j) \delta_i^x \delta_j^y.$$

The counting process and martingale related to the observations on C were defined in Appendix B. We can now use weak convergence of the Kaplan-Meier estimator,

$H_{Tn}^{uc}(t)$ and $H_{Un}^{uc}(u)$, and use the arguments given in Theorem 1 of Datta et al. (2010) to establish the asymptotic normality of \widehat{U}_{1c} .

D: Simulation from Gumbel’s bivariate exponential distributions

As given in Sect. 4, Gumbel’s bivariate exponential distribution has the survival and density functions as

$$\begin{aligned} \bar{W}(x, y) &= e^{-(\lambda_1 x + \lambda_2 y + \delta \lambda_1 \lambda_2 x y)}, \quad x, y > 0, \lambda_1, \lambda_2 > 0, 0 \leq \delta \leq 1, \\ w(x, y) &= W(x, y) \{(1 - \delta)\lambda_1 \lambda_2 + \delta \lambda_1^2 \lambda_2 x + \delta \lambda_1 \lambda_2^2 y + \delta^2 \lambda_1^2 \lambda_2^2 x y\}. \end{aligned}$$

The marginal distribution functions are $F(x) = 1 - e^{-\lambda_1 x}$ and $G(y) = 1 - e^{-\lambda_2 y}$, respectively.

For simulation from the above bivariate distribution, we note that the conditional density and distribution functions of X given y are:

$$\begin{aligned} f(x|y) &= \frac{w(x, y)}{g(y)} = \{(1 - \delta)\lambda_1 + \delta \lambda_1^2 x + \delta \lambda_1 \lambda_2 y + \delta^2 \lambda_1^2 \lambda_2 x y\} e^{-\lambda_1(1 + \delta \lambda_2 y)x} \\ \bar{F}(x|y) &= (1 + \delta \lambda_1 x) e^{-\lambda_1(1 + \delta \lambda_2 y)x}. \end{aligned} \tag{27}$$

We simulate (X, Y) from $W(x, y)$ as follows.

- a. Simulate u from $U(0, 1)$ and define $y = -\log(u)/\lambda_2$. Note that $\bar{G}(Y) \sim U(0, 1)$ and hence, Y has the survival function $\bar{G}(y)$.
- b. For given y simulate x from the conditional distribution $F(x|y)$. This is done by solving the equation $\bar{F}(x|y) = v$, where v is generated from $U(0, 1)$.

E: Jackknife variance and confidence intervals

Our interest is in estimating the unknown variance of a statistic $U(Z)$. Callaert and Veraverbeke (1981) have studied the properties of the jackknife estimator of the variance of a one-sample U-statistic of degree two. The estimator given below coincides with the one obtained by the sample values of the pseudo-values for a U-statistic, see text below Eq. (9) in Callaert and Veraverbeke (1981). Let $\mathbf{z} = (z_i, i = 1, \dots, n)$ be the data from the desired distribution. The jackknife estimator of the variance of $U(Z)$ is obtained by removing one observation at a time and computing the statistic. Let $U_{(-i)}(\mathbf{z}), i = 1, \dots, n$, be the value of the statistic obtained by removing z_i and using only $(n - 1)$ observations. The jackknife variance estimator of $U(Z)$ is

$$\widehat{\text{var}}(U(Z)) = \frac{n - 1}{n} \sum_{i=1}^n (U_{(-i)}(\mathbf{z}) - \bar{U}(\mathbf{z}))^2,$$

where $\bar{U}(\mathbf{z}) = \sum_{i=1}^n U_{(-i)}(\mathbf{z})/n$. Note that $\bar{U}(\mathbf{z})$ is the same as $U(\mathbf{z})$. The confidence interval for $E(U(Z))$ is given by $U(\mathbf{z}) \pm z_{\alpha} \sqrt{\widehat{\text{var}}(U(Z))}$. We refer to Davison and Hinkley (1997) and Efron and Stein (1981) for more details.

For a one-sided test where the null hypothesis is rejected for large values of the observed test statistic, the lower confidence limit of the jackknife confidence interval can be used for testing purpose. The observed data can be taken as an outlier when the null hypothesis is true at α level of significance if $U(\mathbf{z}) - z_{\alpha} \sqrt{\widehat{\text{var}}(U(Z))} > E_{null}(U(Z))$. We employed this technique to obtain the empirical size and power of the tests.

References

- Andersen PK, Borgan Ø, Gill RD, Keiding N (1993) Statistical models based on counting processes. Springer-Verlag, Berlin
- Beare BK, Moon JM (2015) Nonparametric tests of density ratio ordering. *Econ Theory* 31:471–492
- Begun A, Yashin A (2018) Study of the bivariate survival data using frailty models based on Lévy processes. *ASTA Adv Stat Anal*. <https://doi.org/10.1007/s10182-018-0322-y>
- Belzunce F, Martínez-Riquelme C, Mulero J (2016) An introduction to stochastic orders. Elsevier, Amsterdam
- Betensky RA, Finkelstein DM (1999) Estimator for bivariate interval censored data. *Stat Med* 18:3089–3100
- Callaert H, Veraverbeke N (1981) The order of the normal approximation for a studentized U-statistic. *Ann Stat* 9(1):194–200
- Carolan CA, Tebbs JM (2005) Nonparametric tests for and against likelihood ratio ordering in the two-sample problem. *Biometrika* 92:159–171
- Datta S, Bandyopadhyay D, Satten GA (2010) Inverse probability of censoring weighted u-statistics for right-censored data with an application to testing hypotheses. *Scand J Stat* 37:680–700
- Davison AC, Hinkley (1997) Bootstrap methods and their application. Cambridge University Press, D. V
- Diao L, Cook R (2014) Composite likelihood for joint analysis of multiple multistate processes via copulas. *Biostatistics* 15(4):690–705
- de Peres MVO, Achcar JA, Martinez EZ (2020) Bivariate lifetime models in presence of cure fraction: a comparative study with many different copula functions. *Heliyon* 6:e03961
- Dykstra R, Kochar S, Robertson T (1995) Inference for likelihood ratio ordering in the two-sample problem. *J Am Stat Assoc* 90:1034–1040
- Efron B, Stein C (1981) The jackknife estimate of variance. *Ann Stat*. 9(3):586–596
- Frees EW, Carriere JF, Valdez EA (1996) Annuity valuation with dependent mortality. *J Risk Insur* 63(2):229–261
- Geerdens C, Janssen P, Veraverbeke N (2016) Large sample properties of nonparametric copula estimators under bivariate censoring. *Statistics* 50:1036–1055
- Gumbel EJ (1960) Bivariate exponential distributions. *J Am Stat Assoc* 55:698–707
- Herrick D, Mueller H-G, Serachitopal D, Zhang L, Hess KR (2018) mudens: Density estimate, R package version 1.3.2. <https://CRAN.R-project.org/package=mudens>
- Huang H, Zhao Y (2018) Empirical likelihood for the bivariate survival function under univariate censoring. *J Stat Plan Inference* 194:32–46
- Hutson AD (2016) Nonparametric rank based estimation of bivariate densities given censored data conditional on marginal probabilities. *J Stat Distrib Appl* 3:1–14
- Hoeffding W (1948) A class of statistics with asymptotically normal distribution. *Ann Math Stat* 19:293–325
- Jinnah A, Zhao Y (2017) Empirical likelihood inference for the bivariate survival function under univariate censoring. *Commun Stat-Simul Comput* 46:4348–4355
- Kalbfleisch JD, Prentice RL (2002) The statistical analysis of failure time data, 2nd edn. Wiley, New York
- Keilson J, Sumita U (1982) Uniform stochastic ordering and related inequalities. *Can J Stat* 10:181–198
- Lai CD, Xie M (2006) Stochastic ageing and dependence for reliability. Springer, New York
- Lee AJ (2020) U-Statistics: Theory and Practice. CRC Press, Boca Raton

- Lehmann EL (1951) Consistency and unbiasedness of certain nonparametric tests. *Ann Math Stat* 22:165–179
- Lin DY, Ying Z (1993) A simple nonparametric estimator of the bivariate survival function under univariate censoring. *Biometrika* 80:573–581
- Lu X, Burke MD (2008) Nonparametric estimation of linear functionals of a bivariate distribution under univariate censoring. *J Stat Plan Inference* 138:3238–3256
- Luciano E, Spreeuw J, Vigna E (2008) Modelling stochastic mortality for dependent lives. *Insurance* 43:234–244
- Nair NU, Sankaran PG, John P (2018) Modelling bivariate lifetime data using copula. *Metron* 76:133–153
- Nelsen RB (1999) An introduction to copulas. Springer, New York
- Prentice RL, Kalbfleisch JD (2003) Aspects of the analysis of multivariate failure time data. *SORT* 27(1):65–78
- R Core Team (2018) R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna. <https://www.R-project.org/>
- Rondeau V, Mazroui Y, Gonzalez R (2012) frailtypack: an R package for the analysis of correlated survival data with frailty models using penalized likelihood estimation of parametrical estimation. *J Stat Softw* 47(4):1–28
- Roosen J, Hennessy DA (2004) Testing for the monotone likelihood ratio assumption. *J Bus Econ Stat* 22:358–366
- Ross SM (1983) Stochastic processes. Wiley, New York
- Savignoni A, Giard C, Tubert-Bitter P, De Rycke Y (2014) Matching methods to create paired survival data based on an exposure occurring over time: a simulation study with application to breast cancer. *BMC Med Res Methodol* 14:1–18
- Shaked M, Shanthikumar JG (2007) Stochastic orders. Springer, New York
- Shanthikumar JG, David DY (1991) Bivariate characterization of some stochastic order relations. *Adv Appl Probab* 23:642–659
- Shigemizu D, Iwase T, Yoshimoto M, Suzuki Y, Miya F, Borojevich KA, Tsunoda T (2017) The prediction models for postoperative overall survival and disease-free survival in patients with breast cancer. *Cancer Med* 6:1627–1638
- Therneau T (2021) A package for survival analysis in R. R package version 3.2-11. <https://www.CRAN.R-project.org/package=survival>
- Westling T, Downes K J, Small DS (2019) Nonparametric maximum likelihood estimation under a likelihood ratio order. arXiv preprint. <https://arxiv.org/abs/1904.12321>
- Xiong C, ElBarmi H (2002) On detecting change in likelihood ratio ordering. *J Nonparametr Stat* 14:555–568
- Yu T, Li P, Qin J (2017) Density estimation in the two-sample problem with likelihood ratio ordering. *Biometrika* 104:141–152

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.