Assessing the Performance of Predictive Scores: a Unified Copula-based Framework and Algorithm for Numerical Evaluation

Yilong Zhang* Yongzhao Shao†

Abstract

In precision medicine, predictive scores (eg, those from accurate predictive systems or procedures) can be used to inform patients and physicians to make decisions. Survival endpoints are of special importance in precision medicine. Time-dependent AUC and Harrell's C statistic has been routinely used as a global adequacy measure of predictive scores for censored survival outcomes. As new biomarkers and prediction procedures are in rapid development, it is of great interest to develop effective statistical methods and algorithms for comparing predictive power of two or more predictive score systems (eg, incremental AUC under correlated time-dependent ROC curves). Theoretical assessment of correlated predictive scores (eg, hypothesis testing) is complicated for censored time-to-event outcomes. Alternatively, numerical evaluation of the predictive accuracy via comprehensive simulation studies is an attractive approach, however, there is an unmet need to develop effective algorithms to generate multiple correlated predictive scores with given predictive accuracy measure. To fill in this knowledge gap, this paper is to provide a unified copula-based framework for numerically evaluating performance of correlated predictive scores. We designed effective algorithms and an R package to simulate correlated predictive scores with preset accuracy measures such as concordance index or time-dependent AUC for time-to-event outcomes. The simulation based numerical approach is convenient for simultaneously evaluating multiple measures of predictive accuracy with complementary strengths, and also convenient to investigate finite sample properties such as correcting for optimism of a given performance measure using cross-validation or bootstrap.

Keywords: predictive scores, predictive accuracy, c-index, AUC, copula, simulation

1. Introduction

In precision medicine, predictive scores (eg, those from accurate predictive systems or procedures) can be used to inform patients and physicians to make decisions. As a common practice, researchers routinely explore the possibility of adding newly discovered biomarkers into existing regression models as new predictors. The new predictors can be used to obtain new predictive scores that are correlated with the existing one but potentially with higher predictive accuracy. As new biomarkers and prediction procedures are in rapid development, it is of great interest to develop effective statistical methods and algorithms for comparing the predictive power of two or more predictive scores. As is well known, direct theoretical comparison, e.g. via hypothesis testing, on predictive accuracy of such

*Merck Research Laboratories, Rahway, NJ

†Department of Population Health, New York University School of Medicine, New York, NY

correlated scores has been challenging even when the outcome variable Y is binary [Vickers et al., 2011, Demler et al., 2012, Pepe et al., 2013, 2014]. Additionally, survival outcomes play an important role inn precision medicine [Osborn et al., 2015, Scher et al., 2015, Davoli et al., 2017].

Theoretical assessment of the performance of predictive scores for time-to-event outcomes is more complicated due to data censoring. As a specific example, note that Harrell's C statistic has been routinely used as a global adequacy assessment of a predictive score X in predicting corresponding survival time T. Note that Harrell's C statistic $\hat{K}_H(X)$ is commonly regarded as an estimate of the following concordance probability [Pencina and D'Agostino, 2004]

$$K_H(X) = pr(X_1 > X_2 \mid T_1 < T_2), \tag{1}$$

where (T_1, X_1) and (T_2, X_2) are bivariate observations of survival time and predictive score from a pair of randomly selected independent subjects. The concordance index $K_H(X)$ is a useful global assessment of the predictive power of a risk score system X. Let X and Y be two scores or prediction procedures for a subject with event time T. A question of interest is to compare predictive powers between X and Y, i.e., to test the following null hypothesis

$$H_0 : K_H(X) = K_H(Y). \tag{2}$$

Recently, Kang et al. [2015] proposed a formal nonparametric approach using the difference between two Harrell's C statistics, i.e., $\Delta\hat{K} = \hat{K}_H(X) - \hat{K}_H(Y)$ as a test statistic. Harrell's C statistic $\hat{K}_H(X)$ is well known to be a biased estimate of the concordance probability $K_H(X)$ [Uno et al., 2011, Wang and Long, 2016, Han et al., 2017]. In general, $\Delta\hat{K} = \hat{K}_H(X) - \hat{K}_H(Y)$ may not have mean zero under the null hypothesis in formula (2), even for large sample sizes. Given the complexities of direct theoretical assessment, naturally, numerical evaluations of the performance of the proposed method are used. Kang et al. [2015] simulated two predictive scores X and Y from bivariate normal distribution under independent censoring. Under this setting, their simulation results indicated that the $\Delta\hat{K}$ had satisfactory performance in terms of Type I error. Kang *et al.* further provided a publically accessible R package `compareC` to implement their test procedure. Given the nonparametric nature of the proposed method, more extensive simulation studies would be necessary to make sure the proposed nonparametric test has desired type 1 error under various settings beyond the bivariate normal distributions. Unfortunately, neither a rigorous framework nor a general computing algorithm is readily available to simulate correlated predictive scores with given predictive accuracy. This has been an important bottleneck towards effective numerical comparison of predictive scores. To fill in this knowledge gap, this paper is to provide a unified framework for numerically evaluating performance of correlated predictive scores. Utilizing a relationship between copula and commonly used predictive performance measures, we designed effective algorithms and an R package to simulate correlated predictive scores with preassigned accuracy measures such as AUC for binary outcome and concordance index or time-dependent AUC for time-to-event outcomes.

The newly proposed performance evaluation framework for predictive scores and computing algorithms have multiple flexibilities and strengths. First of all, the simulation algorithm is rooted in the theory of copula [Nelsen, 2007]. The flexibility of selecting different families of copula allows investigators to simulate correlated predictive scores with preset AUC or concordance index with more general marginal distributions than the commonly used normal or exponential distributions and various types of correlation structure. Secondly, the same framework can conveniently be used to simulate prediction scores for both binary outcomes as well as time-to-event outcomes. In particular, simulating predictive scores with preassigned time-dependent AUC or given concordance index have never been systematically investigated before. Thirdly, the framework can handle multiple correlated predictive scores corresponding to multiple correlated ROC curves or c-indices of predictive procedures. Last but not least, the numerical evaluation framework is convenient to investigate finite sample properties such as correcting for optimism of a given performance measure using cross-validation or bootstrap [Harrell et al., 1984, Steyerberg et al., 2010].

The paper is organized as follow. In section 2, we first review c-index and time-dependent AUC that are commonly used metrics to assess predictive accuracy in time-to-event data. Secondly, we discuss the relationship between copula, c-index and time-dependent AUC. By using the relationship, we propose the unified algorithm to simulate multiple correlated risk scores with given predictive accuracies. In section 3, we illustrate the proposed algorithm through simulation studies. Section 4 contains discussion and final remarks. Some technical proofs are contained in the Appendix. The proposed algorithms have been implemented in an R package `simuCPP` and available at `https://github.com/elong0527/simuCPP`.

2. Methods

2.1 Predictive Accuracy

Various forms of concordance probability or c-index have been widely used in different clinical areas to assess the overall predictive accuracy [Hanley and McNeil, 1982, Harrell et al., 1984, Gönen and Heller, 2005, Heagerty et al., 2000, Heagerty and Zheng, 2005, Steyerberg et al., 2010, Osborn et al., 2015, Scher et al., 2015, Healy et al., 2016, Zhang and Shao, 2017]. Let T denote the event time and X denote a predictive score or a marker. For two independent copies $\{(X_1, T_1),\ (X_2, T_2)\}$ of (X, T), A concordance probability or c-index is defined as

$$K_H = Pr(X_1 > X_2 \mid T_2 > T_1). \tag{3}$$

Without much of generality, we assume small prediction score X is associated with long survival. Then the range of K_H is between 0.5 and 1. The c-index equals to 1 if a marker has perfect discrimination. When the c-index equal to 0.5, the marker does not have meaningful discriminatory power over the naive prediction based on flipping of a fair coin. Another

concordance probability or k-index is defined as

$$K_G = Pr(T_2 > T_1 \mid X_1 > X_2). \tag{4}$$

Consistent estimates of the k-index exist under the commonly used Cox PH models and transformation models [Gönen and Heller, 2005, Zhang and Shao, 2017]. When T and X are continuous random variables and $Pr(X_1 > X_2) = Pr(T_2 > T_1) = 0.5$, then c-index is equivalent to k-index, i.e. $K_H = K_G$. Therefore, our discussion below on c-index K_H can be directly applied to k-index K_G.

Time-dependent ROC and AUC have also been proposed to assess the overall predictive accuracy for time-to-event outcomes [Heagerty et al., 2000, Heagerty and Zheng, 2005]. Two types of time-dependent AUC were investigated by Heagerty and Zheng [2005]. The first one focuses on *cumulative* "cases" that is defined as

$$\text{AUC}^{\mathbb{C}}(t) = Pr(X_1 > X_2 \mid T_1 \leq t,\ T_2 > t), \tag{5}$$

where $t > 0$ is a cut-point used to define "case" $(T \leq t)$ and "control" $(T > t)$. The second one is for *incident* "case" $(T = t)$ that is defined as

$$\text{AUC}^{\mathbb{I}}(t) = Pr(X_1 > X_2 \mid T_1 = t,\ T_2 > t). \tag{6}$$

The integrated incident time-dependent AUC (iAUC) can also be used to summarize the overall prognostic accuracy defined as

$$\text{iAUC}(\tau) = \int_0^\tau \text{AUC}^{\mathbb{I}}(t) \cdot w(t) dt, \tag{7}$$

where τ is a truncation time and $w(t)$ is a known weight function.

2.2 Copula, concordance probability and time-dependent AUC

The copula of a bivariate random vector (W_1, W_2) can be written as

$$C(u, v) = Pr(W_1 < F_{W_1}^{-1}(u), W_2 < F_{W_2}^{-1}(v)),$$

where F_{W_1} and F_{W_2} are the CDF of W_1 and W_2 respectively. A comprehensive introduction of copula and its properties can be found in [Nelsen, 2007] and summarized in the Appendix. Just as statistical software can generate multivariate normal data, there exist multiple R packages and other computing programs that can generate data with given copulas. These recently developed copula generation algorithms and computing packages are publically available are quite convenient to implement.

Without much loss of generality, we can assume the marginal distributions of the correlated predictive scores and the survival time are given and considered known. These

marginal distributions don't have to be normal or exponential, they can be any distributions. Generating data from a given distribution using the inverse of CDF is straightforward. The main challenge is to generate correlated scores so that the dependence structure between the scores and survival time T would ensure the concordance probability or time-dependent AUC exactly equal to the preset value. The key idea in simulating correlated predicti scores with given concordance probability and time-dependent AUC is to use relationships between various copula families and concordance probability and time-dependent AUC. Dependence between marginal distributions are completely determined by the copula function by Sklar's theorem, and concordance probability and time-dependent AUC are functionaries of the copula. Thus, we can simulate various correlated predictic scores by selecting various copula distributions, and obtain correlated score data with any given concordance probability and time-dependent AUC value by selecting parameters in parametric copula densities. In this subsection, we will provide a few Propositions that characterize the general relationships between concordance probability, time-dependent AUC and copula. These relationships are key in determining the parameters in the various copulas that would ensure the given concordance probability and time-dependent AUC and copula. The basic idea is that, for any given concordance probability and time-dependent AUC, we can identify corresponding parameters and copula families cooresponding to the given values of concordance probability and time-dependent AUC. We can use the copula algorithms and computing packages to simulate data from the copula with given concordance probability and time-dependent AUC. Using the marginal copula data and inverse CDF, we can generate the corrected marginal data with the given marginal distributions and given concordance probability and time-dependent AUC.

Proposition 1. Let $C(u, v)$ denote the copula of a bivariate continuous random variable $(-X, \log(T))$, the concordance probability K in formula (3) can be expressed as

$$K_H = 2 \int \int_{[0,1]^2} C(u, v) dC(u, v). \tag{8}$$

Proposition 1 is a direct corollary of Theorem 5.1.3 in [Nelsen, 2007] by noticing that the Kendall's τ is equal to $2K_H - 1$. Therefore, the c-index is fully determined by the copula of T and X. The concordance probability K_H have an analytical form for most widely used copula family as listed in Table 1.

Copula Name	Parameter	Parameter range	K
Gaussian	ρ	$[-1, 1]$	$\frac{1}{\pi}\arcsin(\rho) + \frac{1}{2}$
Clayton	θ	$\theta > 0$	$\frac{\theta}{2(\theta+2)} + \frac{1}{2}$
Gumbel	θ	$\theta > 1$	$1 - \frac{1}{2\theta}$

Table 1: Concordance probability and copula family

Proposition 2. Let $\pi = \pi_t = Pr(T \leq t)$ and $C(u, v)$ denote the copula of a bivariate continuous random vector $(-X, \log(T))$, and $\partial C(u, v)/\partial v$ exists, then the cumulative and incident time-dependent AUC can be expressed by the copula $C(u, v)$ and $\pi = \pi_t$ as

$$AUC^{\mathbb{C}}(t) = AUC^{\mathbb{C}}(\pi) = \frac{1}{\pi(1-\pi)} \int I(v > \pi) C(u, \pi) dC(u, v). \tag{9}$$

and

$$AUC^{\mathbb{I}}(t) = AUC^{\mathbb{I}}(\pi) = \frac{1}{1-\pi} \int I(u > \pi) \frac{\partial C(u, \pi)}{\partial \pi} dC(u, v) \tag{10}$$

The proof of Proposition 2 is described in Appendix. From Proposition 2, it is clear that both the cumulative and incident time-dependent AUC are fully determined by the copula of T and X and the percentile of time $\pi = Pr(T \leq t)$.

Proposition 3. Under the condition of Proposition 2. Let f_T, S_T and F_T denote the density function, survival function and CDF of T, then the iAUC can be expressed as

$$iAUC(\tau) = iAUC^{\mathbb{I}}(\tau, \pi) = \int \left[\int_{\pi \in [0,1]} \frac{I[\min(v, F_T(\tau)) > \pi] w(F_T^{-1}(\pi))}{(1-\pi) f_T(F_T^{-1}(\pi))} dC(u, \pi) \right] dC(u, v). \tag{11}$$

Proposition 3 is a direct corollary of Proposition 2 by combining formula (7) and (10), and interchange of the order of integration. In general, $iAUC$ depends on the marginal distribution and it is not a desired property for a discrimination metric, because a monotone transformation of T can change the value of $iAUC(\tau)$. A proper selection of the weight function $w(t)$ is important to avoid the dependence on the marginal distribution. Therefore the weight function $w(t)$ should proportional to $f_T(t)(1-\pi)$ and $\int_0^\tau w(t)dt = 1$ that result in $w(t) = 2f_T(t)(1-\pi)/(1-S_T^2(\tau))$. This is the same weight function described by [Heagerty and Zheng, 2005]. With the weight function, iAUC can be written as

$$iAUC(\eta) = \frac{2}{\eta(1+\eta)} \int C(u, \min(v, \eta) dC(u, v), \tag{12}$$

where $\eta = F_T(\tau) = Pr(T < \tau)$ is the percentile of the truncation time τ. When $\tau \to \infty$ or $\eta = 1$, it is clear that $iAUC(1) = K_H$ by comparing formula (8) with (12).

Proposition 4. Under the condition of Proposition 2, let type I truncated c-index be defined as $K_1(\tau) = Pr(X_1 > X_2 \mid T_2 > T_1,\ T_1 < \tau)$ and type II truncated c-index defined as $K_2(\tau) = Pr(X_1 > X_2 \mid T_2 > T_1,\ T_2 < \tau)$, we have

$$iAUC(\tau) = K_1(\tau), \text{ if } w(t) = \frac{2f_T(t)(1-\pi)}{1 - S_T^2(\tau)}, \tag{13}$$

and

$$AUC^{\mathbb{C}}(\tau) = \frac{K_1(\tau)\{1 + S_T(\tau)\} - K_2(\tau)\{1 - S_T(\tau)\}}{2S_T(\tau)}. \tag{14}$$

The proof of Proposition 4 is provided in Appendix. The key is to show that

$$K_1(\eta) = \frac{2}{\eta(1+\eta)} \int C(u, \min(v, \eta) dC(u, v) \tag{15}$$

and

$$K_2(\eta) = \frac{2}{\eta^2} \int I(v < \eta) C(u, v) dC(u, v), \tag{16}$$

where $\eta = F_T(\tau) = pr(T < \tau)$. It is clear that $K_1 = K_2 = K_H$ when $\tau \to \infty$ or $\eta = 1$. The type I truncated c-index has been investigated by Uno et al. [2011] and Wang and Long [2016]. The type II truncated c-index has not been investigated in the literature to our best knowledge. The equivalence of $iAUC$ and K_1 in formula (13) has been discussed in [Heagerty and Zheng, 2005]. The main new results is the connection between cumulative time-dependent AUC and truncated c-index in formula (14).

The proofs of the above Propostions are contained in Appendix. The relationships between the c-index K_H and time-dependent AUCs with parameters in some common copula families are illustrated graphically in Figures 1-3.

2.3 Data generation with pre-defined c-index, time-dependent AUC or iAUC

Propositions 1, 2 and 3 play an essential role for simulating data with pre-defined c-index, time-dependent AUC or $iAUC$. We use c-index as an example and other metrics can be treated similarly. If the propose is to simulate T and X from a Clayton copula with c-index $K = k$. It is easy to find the parameter of Clayton copula $\theta = \theta_0$ from Table 1 or solve the equation (8) numerically for θ. Then data can be generated from Clayton copula with $\theta = \theta_0$. Finally T and X can be simulated by transforming the generated data using pre-defined marginal CDF of $\log(T)$ and $-X$ denoted as F_T and F_X. (e.g. T may follow an exponential distribution and X may follow a normal distribution). Let D denote the censoring time, the independent or conditional independent censoring time can be drawn from a pre-defined distribution F_D or $F_{D|X}$, respectively. The algorithm is summarized as below when censoring time is conditionally independent of event time. Figure 1 illustrates the c-index or equivalently iAUC in formula (12) when $\eta = 1$ as a function of the parameter in Gaussian, Clayton and Gumbel copula.

By using the same algorithm and finding the parameter θ based on the formula for truncated c-index (formula (15) and (16)), time-dependent AUC (formula (9) and (10)) or iAUC (formula (11) and (12)) at a truncation time τ, the same algorithm can be used to simulate data with any given truncated c-index, time-dependent AUC or iAUC based on a predefined copula family.

Algorithm 1 Data generation for known c-index based on a predefined copula family

Set values:

Set the desired concordance probability K_H for predictive score X.

Set a bivariate copula families $C_{XT}(u, v \mid \theta)$.

Set marginal CDF F_X, F_T for $-X$, $\log(T)$.

Set conditional CDF $F_{D|X}$ for censoring time D.

Solve parameter θ from the equation (8) .

Generate random variable (U, V) from copula $C_{XT}(u, v)$.

Transform data by

$$X = -F_X^{-1}(U);\ T = \exp[F_T^{-1}(V)].$$

Generate censoring time D from the distribution $F_{D|X}$ based on X.

2.4 Data generation with multiple correlated c-index, time-dependent AUC or iAUC

The problem of interested in about comparing multiple correlated c-index, time-dependent AUC or iAUC to compare the prognostic accuracy for markers. Let's continue use c-index as an example without loss of generality. An important and unmet question is how to simulate multiple correlated c-index to evaluate proposed hypothesis testing procedures. The existing literatures are either based on multivariate normal distribution or rely on unknown true c-index. A further limitation is that the simulated data are commonly assumed independent censoring.

Our unified data generation framework is based on C-vine copula decomposition that provides a graphical way to illustrate the multivariate random variables via paired copula constructions [Joe, 1996, Aas et al., 2009, Brechmann et al., 2013]. Without loss of generality, we consider two correlated c-index with continuous predictive score X_1 and X_2. Let T denote the event time. Let $C_{X_1X_2T}(u_1, u_2, v_1)$ denote the three dimensional joint copula for $(-X_1, -X_2, \log(T))$. We further define C_{X_1T}, C_{X_2T} and $C_{X_1X_2|T}$ as the paired copula for $(-X_1, \log(T))$, $(-X_2, \log(T))$, and $(-X_1, -X_2)$ conditional on T. By C-Vine copula decomposition, the copula $C_{X_1X_2T}$ can be expressed as

$$C_{X_1X_2T}(u_1, u_2, v) = C_{X_1T}(u_1, v)C_{X_2T}(u_2, v)C_{X_1X_2|T}(u_1, u_2). \tag{17}$$

By Proposition 1, the c-index of X_1 is a function of $C_{X_1T}(u_1, v)$ and the c-index of X_2 is a function of $C_{X_2T}(u_2, v)$. The conditional copula $C_{X_1X_2|T}(u_1, u_2)$ controls the correlation between X_1 and X_2. One key feature of C-Vine copula is that the paired copula $C_{X_1T}(u_1, v)$, $C_{X_2T}(u_2, v)$ and $C_{X_1X_2|T}(u_1, u_2)$ are not necessarily the same. The different structure of paired copula greatly extend the distribution family beyond the multivariate normal distributions.

To illustrate the idea, we consider a scenario for simulating two predictive score and event time with predefined c-index $K_1 = k_1$ and $K_2 = k_2$ by assuming C_{X_1T} is a Gaus-

sian copula,C_{X_2T} is a Clayton copula and $C_{X_1X_2|T}$ is a Gumbel copula. It is easy to find the parameter of Gaussian and Clayton copula from Table 1 or solve the equation (8) numerically.The parameter of $C_{X_1X_2|T}$ can be set differently to represent different correlation between X_1 and X_2. Then the data can be generated based on C-vine copula by using the CDVine R package Brechmann et al. [2013]. Finally X_1, X_2 and T can be simulated by transforming the generated data using pre-defined marginal CDF: F_{X_1}, F_{X_2}, and F_T. The independent or conditional independent censoring time can be drawn from a pre-defined CDF F_D or $F_{D|X_1,X_2}$ respectively. The algorithm is summarized below for simulating two correlated c-index with conditional independent censoring time.

Algorithm 2 Data generation for correlated c-indices based on predefined copula family

Set values:

Set the desired concordance probability K_H for predictive score X_1 and X_2.
Set two bivariate copula families $C_{X_1T}(u_1, v \mid \theta_1)$, $C_{X_2T}(u_2, v \mid \theta_2)$.
Set a known bivariate copula for $C_{X_1,X_2|Z}$.
Set marginal CDF F_{X_1}, F_{X_2}, F_T for $-X_1$, $-X_2$ and $\log(T)$.
Set conditional CDF $F_{D|X_1,X_2}$ for censoring time D.

Solve parameter θ_1 and θ_2 from the equation (8)

Generate random variable (U_1, U_2, V) from copula $C_{X_1X_2T}(u_1, u_2, v)$.

Transform data via

$$X_1 = -F_{X_1}^{-1}(U);\ X_2 = -F_{X_2}^{-1}(U);\ T = \exp[F_T^{-1}(V)].$$

Generate censoring time D from the distribution $F_{D|X_1,X_2}$ based on X_1 and X_2.

In addition to the independent or conditional independent censoring, the dependent censoring are also common in practice. The same framework can be used to simulate multiple correlated c-index with deponent censoring. We added the censoring time D in the C-Vine copula decomposition. Let $C_{X_1X_2DT}$ denote the joint copula for $(-X_1, -X_2, \log(D), \log(T))$. We further denote C_{DT} as the paired copula for (D, T) and other conditional copulas accordingly. Then the C-Vine decomposition can be written as

$$\begin{aligned} C_{X_1X_2DT}(u_1, u_2, w, v) =& C_{X_1T}(u_1, v)C_{X_2T}(u_2, v)C_{DT}(w, v) \\ & C_{X_1D|T}(u_1, w)C_{X_2D|T}(u_2, w)C_{X_1X_2|DT}(u_1, u_2). \end{aligned} \tag{18}$$

For simplicity, if X_1 and X_2 are independent of censoring time D conditional on event time T. The formula (18) can be simplified as

$$C_{X_1X_2DT}(u_1, u_2, w, v) = C_{X_1T}(u_1, v)C_{X_2T}(u_2, v)C_{DT}(w, v)C_{X_1X_2|T}(u_1, u_2). \tag{19}$$

Note that, for independent censoring that D is independent of X_1, X_2, T, the formula (18) is equivalent to the formula (17). The simulation algorithm with dependent censoring

through formula (18) and (19) is similar to the Algorithm 2 by specifying all paired copula and other marginal distributions. Furthermore, the algorithm can be easily extended to multiple correlated c-index beyond two risk scores.

The same framework can be used to generate data for multiple correlated truncated c-index (formula (15) and (16)), time-dependent AUC (formula (9) and (10) or iAUC (formula (11) and (12)) at a truncation time τ.

2.5 Covariates generation for c-index, time-dependent AUC or iAUC

The risk score might not be directly available and require to be estimated by other covariates or markers. Let Z denote a p-dimensional vector, and a risk score $X = g(Z)$ where $g(\cdot)$ is an unknown function that needs to be estimated. For example, linear function that assume $X = g(Z, \beta) = Z\beta$ is widely used, where β is a $p-$dimensional parameter vector. A working model like Cox proportional hazards (PH) model are widely used to estimate β [Uno et al., 2011]. The working model should have a good approximation to the true model and does not necessarily to be the true model.

We considered a parametric transformation $X = g(Z, \beta)$, where the form of the function $g(\cdot, \beta)$ is known e.g. $g(Z, \beta) = Z\beta$. Without loss of generality, we assume the first covariates Z_1 is a continuous random variable. Let $Z_{(-1)}$ denote the rest of covariates excluding the first covariate Z_1. Then we have

$$X = g(Z_1, Z_{(-1)}, \beta). \tag{20}$$

We assume that $Z_1 \to X$ is a one to one map in probability one for any $Z_{(-1)}$ and β. For illustration, we consider the linear function $X = Z\beta = Z_1\beta_1 + Z_{(-1)}\beta_{(-1)}$, where $\beta_1 \neq 0$ and $\beta_{(-1)}$ are the parameter vector after removing β_1. Then we have

$$Z_1 = \frac{1}{\beta}[X - Z_{(-1)}\beta_{(-1)}]. \tag{21}$$

To generate the covariates for a risk score, we can first simulated risk score X with desired predictive metric. Secondly $Z_{(-1)}$ can be simulated from a pre-defined distribution. After we set the parameter β, we can create Z_1 by using the formula (21). The algorithm is summarized as below by using c-index as an example. The same framework as in Algorithm 2 can be used to generate multiple correlated c-index with covariates.

3. Numerical Illustration

In this section, we conducted numerical studies to illustrate the performance of the proposed data generation framework for AUC and c-index. Let $\pi = Pr(T < t_{0.5}) = 0.5$ and $Y = I(T < t_{0.5})$. We illustrate the Algorithm 2 to simulate binary outcome Y with known AUC. HereX_1 and X_2 are two correlated risk scores. Let copula of $(-X_1, \log T)$ be a Gaussian

Algorithm 3 Data generation for given c-index with multiple covariates based on pre-defined copula family

Set values:
Set the desired concordance probability K_H for predictive score X.
Set the function $X = g(Z, \beta)$ with known parameter β.
Set the distribution of $Z_{(-1)}$.
Generate X, T and D based on the algorithm 1.
Generate $Z_{(-1)}$ from the defined distribution.
Generate Z_1 by solving equation (21).

copula with parameter ρ and the copula of $(-X_2, \log T)$ a Clayton copula with parameter θ. The conditional copula $(-X_1, -X_2) \mid T$ is a Gumbel copula with parameter ϕ. To achieve desired $AUC = 0.75$, the parameter $\rho = 0.54$ and $\theta = 1.16$ were solved numerically by using the equation (9) when $\pi = 0.5$ for Gaussian and Clayton copulas. The data was simulated by using the R package `simuCPP` that implement the proposed algorithms in this paper with $N = 100$ samples and ϕ =1, 1.5, 3, 5, and 10. Finally, we estimate the AUC empirically by

$$\widehat{AUC} = \frac{\sum_{i \neq j} I(X_i > X_j, Y_i = 1, Y_j = 0)}{\sum_{i \neq j} I(Y_i = 1, Y_j = 0)}.$$

The whole procedure was repeated $1,000$ times and the results were summarized in Table (2). From Table (2), it is clear that the estimators $\widehat{AUC}$ for both X_1 and X_2 equal to 0.75. The Kendall's concordance correlation between X_1 and X_2 is controlled by parameter ϕ in the conditional copula $(X_1, X_2) \mid T$. Specifically, a larger value of ϕ provides a higher value of Kendall's concordance correlation between X_1 and X_2. It is worth to note that the simulated AUC is equivalent to the cumulative time dependent AUC in formula (5) with a cut-point at $t_{0.5}$.

ϕ	$\widehat{AUC}_{X_1}(sd)$	$\widehat{AUC}_{X_2}(sd)$	Kendall's correlation between X_1 and $X_2(sd)$
1	0.75(0.05)	0.75(0.05)	0.19(0.07)
1.5	0.75(0.05)	0.75(0.05)	0.44(0.06)
3	0.75(0.05)	0.75(0.05)	0.69(0.04)
5	0.75(0.05)	0.75(0.05)	0.79(0.02)
10	0.75(0.05)	0.75(0.05)	0.85(0.02)

Table 2: Simulation results for correlated scores X_1, X_2 such that $AUC = 0.75$

We further illustrate the Algorithm 2 to simulate time-to-event outcome T with known c-index. Similarly, let the copula of $(-X_1, \log(T))$ be a Gaussian copula with parameter ρ and the copula of $(-X_2, \log(T))$ a Clayton copula with parameter θ. The conditional copula $(-X_1, -X_2) \mid T$ is a Gumbel copula with parameter ϕ. To achieve the desired concordance

probability $K_H = 0.75$, Table (1) infers that $\rho = 0.71$ and $\theta = 2$ for Gaussian and Clayton copulas. Data were generated with $N = 100$ samples and ϕ =1, 1.5, 3, 5, and 10. The c-index can be empirically estimated by

$$\hat{K}_H = \frac{\sum_{i \neq j} I(X_i < X_j, T_i > T_j)}{\sum_{i \neq j} I(X_i < X_j)}.$$

The whole procedure was repeated 1,000 times and the results are summarized in Table 3. From Table 3, it is clear that the estimators $\hat{K}_H$ for both X_1 and X_2 equal to 0.75 in all scenarios. The Kendall's concordance correlation between X_1 and X_2 are controlled by parameter ϕ in the conditional copula $(-X_1, -X_2) \mid T$. Specifically, the larger value of ϕ provides a higher value of Kendall's concordance correlation between X_1 and X_2.

ϕ	$\hat{K}_H(X_1)(sd)$	$\hat{K}_H(X_2)(sd)$	Kendall's correlation between X_1 and X_2 (sd)
1	0.75(0.03)	0.75(0.03)	0.33(0.06)
1.5	0.75(0.02)	0.75(0.03)	0.53(0.05)
3	0.75(0.02)	0.75(0.03)	0.73(0.03)
5	0.75(0.02)	0.75(0.03)	0.80(0.02)
10	0.75(0.03)	0.75(0.03)	0.85(0.02)

Table 3: Simulation results for correlated scores X_1, X_2 such that $K_H = 0.75$.

4. Figures

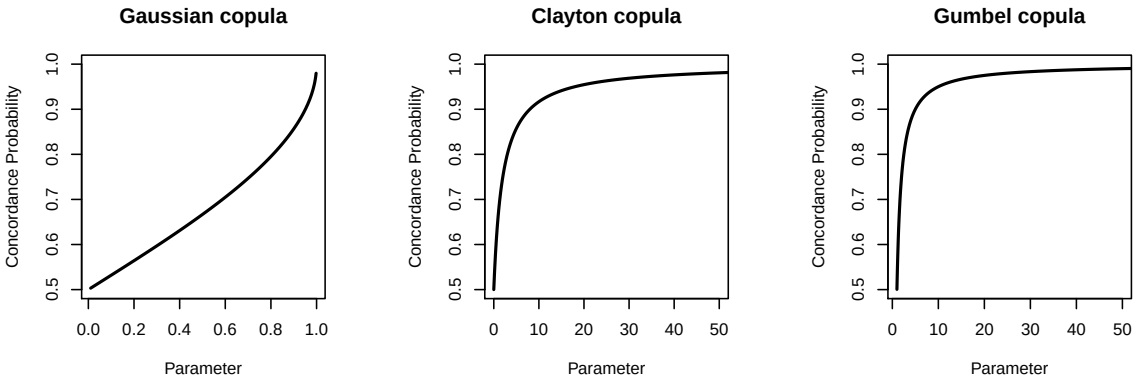

Figure 1: The relationship between K_H and copula parameter in Gaussian, Clayton and Gumbel copulas

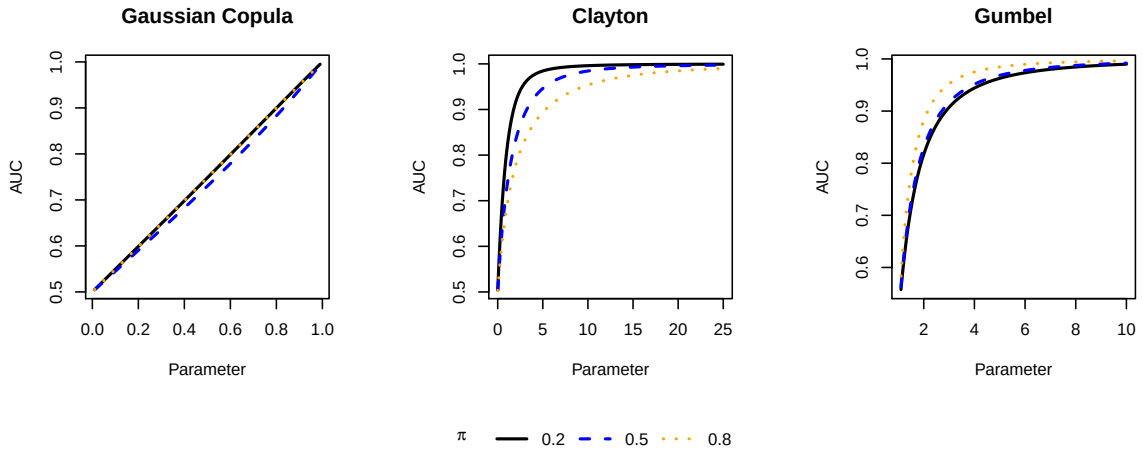

Figure 2: The relationship between cumulative time-dependent AUC, $AUC^{\mathbb{C}}(\pi)$ and copula parameter in Gaussian, Clayton and Gumbel copulas, when π equal to 0.2, 0.5 and 0.8

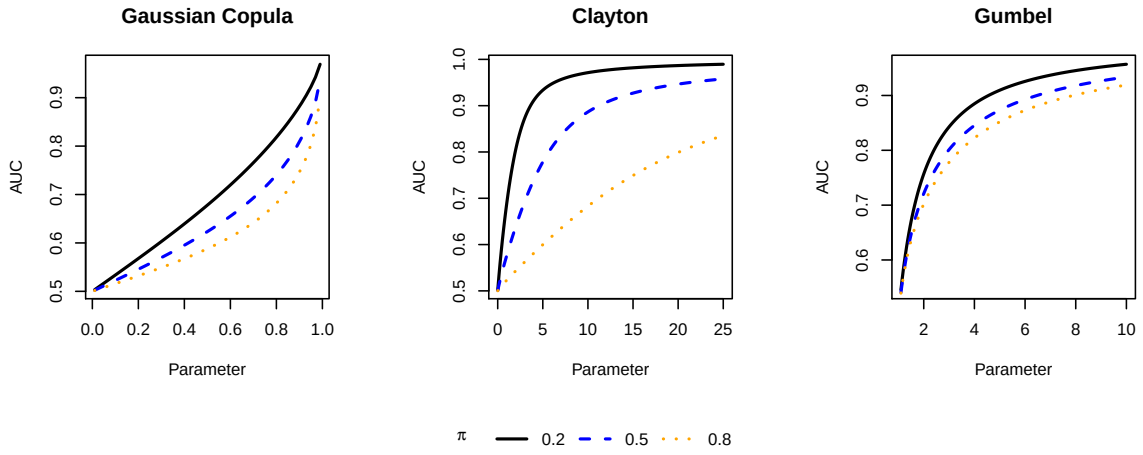

Figure 3: The relationship between incident time-dependent AUC, $AUC^{\mathbb{I}}(\pi)$ and copula parameter in Gaussian, Clayton and Gumbel copulas, when π equal to 0.2, 0.5 and 0.8

5. Discussion

In this paper, we proposed a unified copula based framework to generate data for multiple correlated risk scores with pre-defined predictive metrics. By using the relationship

between predictive metrics and copula, the unified simulation framework and algorithm have multiple benefits. 1) The algorithm starts from pre-assigned accuracy value. 2) The algorithm can conveniently control the correlation between risk scores. 3) The algorithm is flexible by choosing different kinds of copula to reflect diverse association structure. 4) The algorithm allows the risk score to be a function of other covariates. 5) The algorithm allows different kinds of censoring type including independent censoring, conditional independent consoring and dependent censoring. As shown in Han et al. [2017], making conclusions on non-parametric methods using numerical evaluation based only on multivariate normal distribution could result in biased conclusion. The proposed general framework largely extend the distribution family beyond multivariate normal distribution to evaluate non-parametric or semi-parametric estimation and hypothesis testing method for predictive metrics.

The framework works both for time-to-event outcomes and binary outcomes as illustrated in Section 3. Even through the focous of this paper is on c-index, time-dependent AUC and iAUC, the same framework can be applied to other predictive metrics as long as a relationship between the metric and copula can be constructed. Furthermore, we show that the copula theory is of fundamental importance to investigate predictive metrics. The relationship between predictive metrics can be verified by using copula. The area deserves further systematic investigation.

6. Acknowledgements

This research was supported in part by NIH grants P30 AG0851, 2P30 CA16087 and 5P30 ES00260.

References

Kjersti Aas, Claudia Czado, Arnoldo Frigessi, and Henrik Bakken. Pair-copula constructions of multiple dependence. *Insurance: Mathematics and economics*, 44(2):182–198, 2009.

Eike Christian Brechmann, Ulf Schepsmeier, et al. Modeling dependence with c-and d-vine copulas: The r-package cdvine. *Journal of Statistical Software*, 52(3):1–27, 2013.

David G Clayton. A model for association in bivariate life tables and its application in epidemiological studies of familial tendency in chronic disease incidence. *Biometrika*, 65 (1):141–151, 1978.

Teresa Davoli, Hajime Uno, Eric C Wooten, and Stephen J Elledge. Tumor aneuploidy correlates with markers of immune evasion and with reduced response to immunotherapy. *Science*, 355(6322):eaaf8399, 2017.

Olga V Demler, Michael J Pencina, and Ralph B D'Agostino. Misuse of delong test to compare aucs for nested models. *Statistics in medicine*, 31(23):2577–2587, 2012.

Mithat Gönen and Glenn Heller. Concordance probability and discriminatory power in proportional hazards regression. *Biometrika*, 92(4):965–970, 2005.

Emil J Gumbel. Bivariate exponential distributions. *Journal of the American Statistical Association*, 55(292):698–707, 1960.

Xiaoxia Han, Yilong Zhang, and Yongzhao Shao. On comparing 2 correlated c indices with censored survival data. *Statistics in Medicine*, 2017.

James A Hanley and Barbara J McNeil. The meaning and use of the area under a receiver operating characteristic (roc) curve. *Radiology*, 143(1):29–36, 1982.

Frank E Harrell, Kerry L Lee, Robert M Califf, David B Pryor, and Robert A Rosati. Regression modelling strategies for improved prognostic prediction. *Statistics in medicine*, 3(2):143–152, 1984.

Patrick J Heagerty and Yingye Zheng. Survival model predictive accuracy and roc curves. *Biometrics*, 61(1):92–105, 2005.

Patrick J Heagerty, Thomas Lumley, and Margaret S Pepe. Time-dependent roc curves for censored survival data and a diagnostic marker. *Biometrics*, 56(2):337–344, 2000.

Mark A Healy, Evan Reynolds, Mousumi Banerjee, and Sandra L Wong. Lymph node ratio is less prognostic in melanoma when minimum node retrieval thresholds are not met. *Annals of Surgical Oncology*, pages 1–7, 2016.

Harry Joe. Families of m-variate distributions with given margins and m (m-1)/2 bivariate dependence parameters. *Lecture Notes-Monograph Series*, pages 120–141, 1996.

Le Kang, Weijie Chen, Nicholas A Petrick, and Brandon D Gallas. Comparing two correlated c indices with right-censored survival outcome: a one-shot nonparametric approach. *Statistics in medicine*, 34(4):685–703, 2015.

Roger B Nelsen. *An introduction to copulas*. Springer Science & Business Media, 2007.

David Oakes. A model for association in bivariate survival data. *Journal of the Royal Statistical Society. Series B (Methodological)*, pages 414–422, 1982.

David PJ Osborn, Sarah Hardoon, Rumana Z Omar, Richard IG Holt, Michael King, John Larsen, Louise Marston, Richard W Morris, Irwin Nazareth, Kate Walters, et al. Cardiovascular risk prediction models for people with severe mental illness: results from the prediction and management of cardiovascular risk in people with severe mental illnesses (primrose) research program. *JAMA psychiatry*, 72(2):143–151, 2015.

Michael J Pencina and Ralph B D'Agostino. Overall c as a measure of discrimination in survival analysis: model specific population value and confidence interval estimation. *Statistics in medicine*, 23(13):2109–2123, 2004.

Margaret S Pepe, Holly Janes, and Christopher I Li. Net risk reclassification p values: valid or misleading? *Journal of the National Cancer Institute*, 106(4):dju041, 2014.

Margaret Sullivan Pepe, Kathleen F Kerr, Gary Longton, and Zheyu Wang. Testing for improvement in prediction model performance. *Statistics in medicine*, 32(9):1467–1482, 2013.

Howard I Scher, Glenn Heller, Arturo Molina, Gerhardt Attard, Daniel C Danila, Xiaoyu Jia, Weimin Peng, Shahneen K Sandhu, David Olmos, Ruth Riisnaes, et al. Circulating tumor cell biomarker panel as an individual-level surrogate for survival in metastatic castration-resistant prostate cancer. *Journal of Clinical Oncology*, 33(12):1348–1355, 2015.

M Sklar. Fonctions de repartition an dimensions et leurs marges. *Publ. Inst. Statist. Univ. Paris*, 8:229–231, 1959.

Ewout W Steyerberg, Andrew J Vickers, Nancy R Cook, Thomas Gerds, Mithat Gonen, Nancy Obuchowski, Michael J Pencina, and Michael W Kattan. Assessing the performance of prediction models: a framework for some traditional and novel measures. *Epidemiology (Cambridge, Mass.)*, 21(1):128, 2010.

Hajime Uno, Tianxi Cai, Michael J Pencina, Ralph B D'Agostino, and LJ Wei. On the c-statistics for evaluating overall adequacy of risk prediction procedures with censored survival data. *Statistics in medicine*, 30(10):1105–1117, 2011.

Andrew J Vickers, Angel M Cronin, and Colin B Begg. One statistical test is sufficient for assessing new predictive markers. *BMC medical research methodology*, 11(1):13, 2011.

Ming Wang and Qi Long. Addressing issues associated with evaluating prediction models for survival endpoints based on the concordance statistic. *Biometrics*, 2016.

Yilong Zhang and Yongzhao Shao. Concordance measure and discriminatory accuracy in transformation cure models. *Biostatistics*, 2017(1):kxx016, 2017.

7. Appendix

7.1 Introduction to Copulas

For any bivariate random variables $X = (X_1, X_2)$ with a joint cumulative distribution function (CDF) $F(x_1, x_2)$, Sklar's theorem [Sklar, 1959] ensures that every bivariate CDF can be written as

$$F(x_1, x_2) = C(F_1(x_1), F_2(x_2)),$$

where F_1, F_2 are marginal CDF of X_1 and X_2, respectively, and the function $C(\cdot, \cdot)$ a bivariate probability distribution with uniform marginal probability distributions, called a copula. We provide a short introduction to bivariate copulas for completeness. A comprehensive introduction to copula and its properties can be found in Nelsen [2007].

Definition: A bivariate copula is a bivariate function $C(\cdot, \cdot)$ whose domain is $[0, 1]^2$ with the following properties:

1. For every u, v in $[0, 1]$,
$$C(u, 0) = 0 = C(0, v)$$
and
$$C(u, 1) = u \text{ and } C(1, v) = v;$$
2. For every u_1, u_2, v_1, v_2 in $[0, 1]$ such that $u_1 \leq u_2$ and $v_1 \leq v_2$,
$$C(u_2, v_2) - C(u_2, v_1) - C(u_1, v_2) + C(u_1, v_1) \geq 0.$$

There are several important and widely investigated copula families. If two continuous random variables are independent, they have product copula defined as

$$C(u, v) = uv. \tag{22}$$

If two continuous random variables (X_1, X_2) follow a bivariate normal distribution with correlation coefficient ρ, they have the Gaussian copula:

$$C(u, v) = \frac{1}{2\pi\sqrt{1-\rho^2}} \int_{-\infty}^{\Phi^{-1}(u)} \int_{-\infty}^{\Phi^{-1}(v)} \exp\left[\frac{-(s^2 - 2\rho st + t^2)}{2(1-\rho^2)}\right] dsdt, \tag{23}$$

(with $\rho \neq -1$, or 1) where Φ is the CDF of the standard normal distribution. Here $\rho = \text{cor}(X_1, X_2)$ is the Pearson correlation between X_1 and X_2.

There is another important class of copula called the Archimedean copula that includes several important copula families [Nelsen, 2007]. With a continuous, strictly decreasing generating function $\varphi(\cdot)$ from $[0, 1]$ to $[0, \infty]$ such that $\varphi(1) = 0$, an Archimedean copula has the form

$$C(u, v) = \varphi^{-1}(\varphi(u) + \varphi(v)). \tag{24}$$

Theorem 4.1.4 in Nelsen [2007] ensures that formula (24) is a copula if and only if φ is convex. The Archimedean copula is commonly used in practice mainly for three reasons: (1) Archimedean copulas are easy to construct by specifying the generating function. (2) There are a large variety of copula families that belong to the Archimedean copula class. (3) the Archimedean copula has many nice properties as discussed in Nelsen [2007] that can help researchers to investigate the theoretical properties.

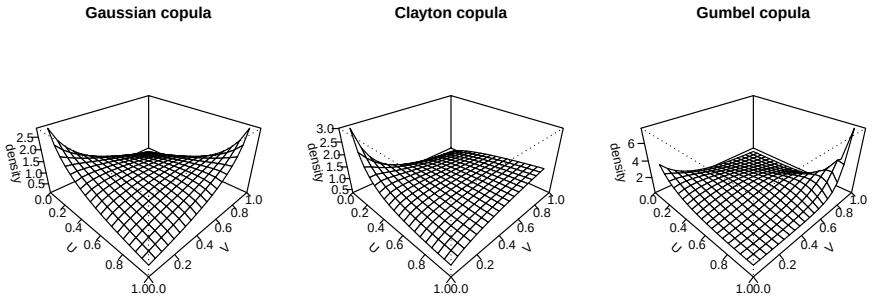

Figure 4: The density of Gaussian ($\rho = 0.5$), Clayton ($\theta = 0.5$) and Gumbel ($\theta = 2$) copulas

Important Archimedean copulas include Clayton copula [Clayton, 1978, Oakes, 1982], Gumbel copula [Gumbel, 1960]. Both of Clayton and Gumbel copula are extended from bivariate exponential distributions. Clayton copula has generating function $\varphi_\theta(t) = (t^{-\theta} - 1)/\theta$ and defined as

$$C(u, v) = \left[\max\left(u^{-\theta} + v^{-\theta} - 1, 0\right)\right]^{-1/\theta}, \tag{25}$$

where θ is in $[-1, \infty)$ except 0. The relationship between Clayton copula parameter θ and Kendall's τ is $\tau = \theta/(\theta + 2)$.

Gumbel copula has generating function $\varphi_\theta(t) = (-\log t)^\theta$ and defined as

$$C(u, v) = \exp\left[-\left((-\log(u))^\theta + (-\log(v)^\theta\right)^{1/\theta}\right] \tag{26}$$

where θ is in $[1, \infty)$. The relationship between Gumbel copula parameter θ and Kendall's τ is $\tau = 1 - 1/\theta$. In Figure 4, we show the density functions of Gaussian, Clayton and Gumbel copulas.

7.2 Proof of propositions

We prove the Proposition 2 and 4 in this appendix.

Proof. of Proposition 2. We first consider cumulative time-dependent AUC. Let $p_{XT} = Pr(X_1 > X_2, T_1 \leq t, T_1 > t)$, where t is a cut-point. By Bayes theorem and independence of two observations, we have

$$AUC = \frac{p_{XT}}{\pi(1-\pi)}, \tag{27}$$

where $\pi = Pr(T \leq t)$. Let F_X denote the CDF of $-X$ and F_T denote the CDF of $\log(T)$. Let $U_1 = F_X(-X_1)$ and $U_2 = F_X(-X_2)$, $V_1 = F_T(\log(T_1))$ and $V_2 = F_T(\log(T_2))$, we have

$$\pi = Pr(T \leq t) = F_T(\log(t))$$

By the condition of theorem and we have

$$\begin{aligned} p_{XY} &= pr(U_2 > U_1, V_1 \leq F_T(\log(t)), V_2 > F_T(\log(t))) \\ &= pr(U_2 > U_1, V_1 \leq \pi, V_2 > \pi). \end{aligned}$$

Notice that the copula of (X_1, T_1) and (X_2, T_2) are

$$C_{XT}(u, v) = pr(U_1 < u, V_1 < v) = pr(U_2 < u, V_2 < v)$$

We have

$$\begin{aligned} p_{XT} &= \int I(v > \pi) pr(U_1 < u, V_1 < \pi) dC_{XT}(u, v) \\ &= \int I(v > \pi) C_{XT}(u, \pi) dC_{XT}(u, v). \end{aligned} \tag{28}$$

By formula (27) we get the conclusion for cumulative time-dependent AUC as

$$AUC^{\mathbb{C}}(t) = AUC^{\mathbb{C}}(\pi) = \frac{1}{\pi(1-\pi)} \int I(v > \pi) C(u, \pi) dC(u, v). \tag{29}$$

Similarly, we consider $AUC^{\mathbb{I}}(t)$ as below.

$$\begin{aligned} AUC^{\mathbb{I}}(t) &= Pr(X_1 > X_2 \mid T_1 = t, T_2 > t) \\ &= Pr(U_2 > U_1 \mid V_1 = \pi, V_2 > \pi) \\ &= \lim_{\delta \to 0} \frac{Pr(U_2 > U_1, V_1 \in (\pi, \pi + \delta), V_2 > \pi)}{Pr(V_1 \in (\pi, \pi + \delta)) Pr(V_2 < \pi)} \\ &= \lim_{\delta \to 0} \frac{1}{(1-\pi)\delta} \int I(v > \pi) Pr(U_1 < u, V_1 \in (\pi, \pi + \delta)) dC_{XT}(u, v) \\ &= \frac{1}{(1-\pi)} \int I(v > \pi) \lim_{\delta \to 0} \frac{C_{XT}(u, \pi + \delta) - C_{XT}(u, \pi)}{\delta} dC_{XT}(u, v) \\ &= \frac{1}{1-\pi} \int I(v > \pi) \frac{\partial C(u, \pi)}{\partial \pi} dC(u, v). \end{aligned}$$

Therefore, we have

$$AUC^{\mathbb{I}}(t) = AUC^{\mathbb{I}}(\pi) = \frac{1}{1-\pi}\int I(u > \pi)\frac{\partial C(u,\pi)}{\partial \pi} dC(u,v).$$

□

Proof. of Proposition 4. We first show formula (15) and (16) hold. Let F_X denote the CDF of $-X$ and F_T denote the CDF of $\log(T)$. Let $U_1 = F_X(-X_1)$ and $U_2 = F_X(-X_2)$, $V_1 = F_T(\log(T_1))$ and $V_2 = F_T(\log(T_2))$. We first consider K_1.

$$Pr(X_1 > X_2 \mid T_2 > T_1, T_1 < \tau) = \frac{Pr(U_2 > U_1, V_2 > V_1, V_1 < \eta)}{Pr(V_2 > V_1, V_1 < \eta)} \tag{30}$$

Notice that

$$\begin{aligned} Pr(V_2 > V_1, U_2 > U_1, U_1 < \eta) &= \int Pr(U_1 < u, V_1 < \min(v,\eta)) dC_{XT}(u,v) \\ &= \int C_{XT}(u, \min(v,\eta)) dC_{XT}(u,v), \end{aligned} \tag{31}$$

and

$$Pr(V_2 > V_1, V_1 < \eta) = \frac{1}{2}(\eta - \eta^2). \tag{32}$$

Combining equation (30), (31) and (32) we have

$$K_1 = \frac{2}{1-(1-\eta)^2}\int C_{XT}(u, \min(v,\eta)) dC_{XT}(u,v).$$

Similarly we consider $K_2(\tau)$.

$$\begin{aligned} Pr(X_1 > X_2 \mid T_2 > T_1, T_2 < \eta) &= Pr(U_2 > U_1 \mid V_2 > V_1, V_2 < \eta) \\ &= \frac{Pr(U_2 > U_1, V_2 > V_1, V_2 < \eta)}{Pr(V_2 > V_1, V_2 < \tau)} \end{aligned} \tag{33}$$

Notice that

$$\begin{aligned} Pr(U_2 > U_1, V_2 > V_1, V_2 < \eta) &= \int I(v < \eta) Pr(U_1 < u, V_1 < v) dC_{XT}(u,v) \\ &= \int I(v < \eta) C_{XT}(u,v) dC_{XT}(u,v). \end{aligned} \tag{34}$$

and

$$\begin{aligned} Pr(V_2 > V_1, V_2 < \eta) &= \int I(v < \eta) pr(V_1 < v) dv \\ &= \int I(v < \eta) v dv \\ &= \frac{1}{2}\eta^2 \end{aligned} \tag{35}$$

Combining equation (33), (34) and (35) we have

$$K_2(\tau) = \frac{2}{\eta^2}\int I(v<\eta)C_{XT}(u,v)dC_{XT}(u,v).$$

We show formula (13) hold by insert the weight function $w(t) = 2f_T(t)(1-\pi)/(1-S_T^2(\tau))$ into formula (11) and notice that $S_T(t) = pr(T>t) = 1-\eta$ by definition. So

$$\begin{aligned} iAUC(\tau) &= \int\left[\int_{\pi\in[0,1]}\frac{2I[\min(v,F_T(\tau))>\pi]}{1-S_T^2(\tau)}dC(u,\pi)\right]dC(u,v) \\ &= \frac{2}{1-(1-\eta)^2}\int C_{XT}(u,\min(v,F_T(\tau))dC_{XT}(u,v) \\ &= K_1(\tau). \end{aligned}$$

Now we prove the formula (14). Let $p_{XT} = Pr(X_1 > X_2, T_1 \le t, T_2 > t)$, as shown in formula (28) we have

$$p_{XT}(\pi) = \int I(v>\pi)C_{XT}(u,\pi)dC_{XT}(u,v)$$

Then we have

$$\frac{\partial p_{XT}(\pi)}{\partial\pi} = \int I(v>\pi)\frac{\partial C_{XT}(u,\pi)}{\partial\pi}dC_{XT}(u,v) - \int_u C(u,\pi)dC(u,\pi). \tag{36}$$

By taking integration at two sides in $(0,\eta)$, we have

$$\begin{aligned} p_{XT}(\eta) &= \int\int_\pi I(\pi<\min(v,\eta))\frac{\partial C_{XT}(u,\pi)}{\partial\pi}d\pi dC_{XT}(u,v) - \\ &\quad \int_{u,v} I(\pi<\eta)C_{XT}(u,\pi)dC_{XT}(u,\pi). \\ &= \int C_{XT}(u,\min(v,\eta))dC_{XT}(u,v) - \int_{u,v} I(\pi<\eta)C_{XT}(u,\pi)dC_{XT}(u,\pi) \end{aligned}$$

By formula (27) and (29), we have

$$p_{XT}(\eta) = AUC^{\mathbb{C}}(\eta)\eta(1-\eta).$$

Combining formula (15) and (16) we have

$$AUC^{\mathbb{C}}(\eta)\eta(1-\eta) = \frac{1-(1-\eta)^2}{2}K_1(\eta) - \frac{\eta^2}{2}K_2(\eta).$$

By definition we have $S_T(t) = pr(T>t) = 1-\eta$, we get the conclusion that

$$AUC^{\mathbb{C}}(\tau) = \frac{K_1(\tau)[1+S_T(\tau)] - K_2(t)[1-S_T(t)]}{2S_T(\tau)}$$

□