HACSurv: A Hierarchical Copula-Based Approach for Survival Analysis with Dependent Competing Risks

Xin Liu[1,2], Weijia Zhang[3], Min-Ling Zhang[1,2]

[1]Southeast University, [2]Key Laboratory of CNI, [3]The University of Newcastle

Highlights

- HACSurv is a data-driven method to model dependencies among competing risks and censoring.
- Modeling dependencies leads to better survival outcome prediction.

Motivation

- Existing methods typically assume independence among competing risks and censoring.
- In reality, competing risks may be dependent, e.g., as breast cancer progresses, the risk of cardiovascular disease may increase.
- It is essential to capture and model such dependencies for accurate survival analysis.

Key Idea

- Use a hierarchical Archimedean copula (HAC) parameterized by neural network to model the dependency structure.
- Use monotonic neural density estimators (MONDE) to model marginal survival distributions.
- Train the model on right-censored survival data via maximum likelihood estimation.

Problem Setup

Each sample in the survival dataset $D = \{(\boldsymbol{x_i}, t_i, e_i)\}_{i=1}^{N}$ contains a covariate vector $\boldsymbol{x_i}$, an observed time t_i, and an event indicator $e_i \in \{0,1, \ldots, K\}$, where $e_i = 0$ denotes censoring.

In the competing risks setting, only one event is observed for each patient, and t_i is the minimum of all possible event and censoring time.

Survival Likelihood

- Initial form (for each sample):

$$\mathcal{L} = \prod_{k=0}^{K} [\Pr(T_k = t, \{T_i > t\}_{i \neq k} | \boldsymbol{x})]^{\mathbb{1}\{e=k\}}$$

- Survival likelihood is easy to estimate with independent assumption (we omit $\boldsymbol{x}$):

$$\mathcal{L}_{\text{Indep}} = \prod_{k=0}^{K} \left[f_k(t) \prod_{i \neq k} S_i(t) \right]^{\mathbb{1}\{e=k\}}$$

- We use a copula to model the joint survival distribution:

$$\mathcal{L} = \prod_{k=0}^{K} \left\{ f_{T_k}(t) \frac{\partial}{\partial u_k} C(u_0, \ldots, u_K) \right\}^{\mathbb{1}\{e=k\}}$$

Copula, AC and HAC

- A copula is a probability distribution with uniform margins. C: $[[0,1]^n \rightarrow [0,1]$. According to Sklar's Theorem [1], any joint survival distribution can be modeled by a copula with corresponding marginal survival distributions:

$$\Pr(T_0 > t, \ldots, T_K > t \mid \boldsymbol{x}) = C\left(S_{T_0|X}(t|\boldsymbol{x}), \ldots, S_{T_K|X}(t|\boldsymbol{x})\right)$$

- Archimedean copulas (AC) is a widely used family of copulas that can represent different tail dependencies using **a one-dimensional generator**:

$$C(\mathbf{u}) = \varphi\left(\varphi^{-1}(u_1) + \cdots + \varphi^{-1}(u_d)\right)$$

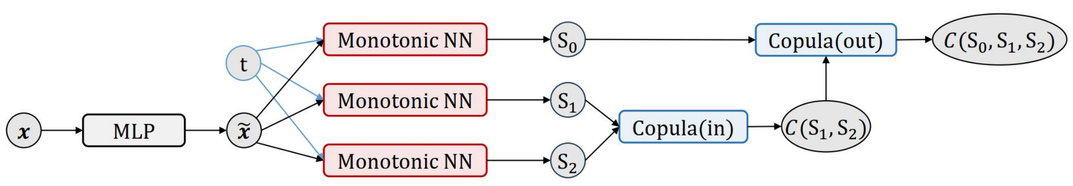
- Archimedean copulas can not model **asymmetric** dependency structures. In dimensions higher than 2, dependencies are rarely symmetric. HAC is used to model asymmetric dependencies. Two competing risks example:

$$C(u_1, u_2, u_3) = \varphi_0\left(\varphi_0^{-1}(u_1) + \varphi_0^{-1} \circ \varphi_1\left(\varphi_1^{-1}(u_2) + \varphi_1^{-1}(u_3)\right)\right)$$

- For the survival likelihood, when $k = 1$:

$$\frac{\partial}{\partial u_1} C(u_0, u_1, u_2) = \left. \frac{\partial C(u_0, u_1, u_2)}{\partial C_{in}(u_1, u_2)} \cdot \frac{\partial C_{in}(u_1, u_2)}{\partial u_1} \right|_{\substack{u_0 = S_{T_0}(t) \\ u_1 = S_{T_1}(t) \\ u_2 = S_{T_2}(t)}}$$

Overview of HACSurv



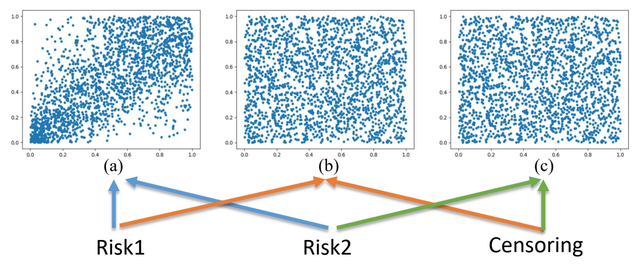
We use the generative Archimedean copula proposed by Ng *et al.* [2] to model AC and HAC.

Determine the Structure of HAC

- Capture the copulas between each pair of competing risks and censoring. Similar to DCSurvival [3], we compute pairwise likelihood:

$$\mathcal{L}_{k_1,k_2} = \left[f_{T_{k_1}}(t) \cdot \frac{\partial}{\partial u_{k_1}} C(u_{k_1}, u_{k_2}) \Big|_{\substack{u_{k_1} = S_{T_{k_1}}(t) \\ u_{k_2} = S_{T_{k_2}}(t)}} \right]^{\mathbb{1}\{e=k_1\}} \cdot \left[f_{T_{k_2}}(t) \cdot \frac{\partial}{\partial u_{k_2}} C(u_{k_1}, u_{k_2}) \Big|_{\substack{u_{k_1} = S_{T_{k_1}}(t) \\ u_{k_2} = S_{T_{k_2}}(t)}} \right]^{\mathbb{1}\{e=k_2\}}$$

- For 2 risks scenario, we should get 3 pairwise copulas. One constraint in determining the HAC structure is that the inner copulas must exhibit stronger dependency than the outer copulas.



Training inner generator with Re-generation Trick

- The copula obtained in the previous stage can be directly used as the outer copula, but we need to learn the parameters of the inner copulas.
- Re-generation Trick:
  - We sample 2-dimensional data from the AC generator obtained in the previous stage.
  - We use this new data to construct a dataset and we train the inner generator via maximum likelihood estimation.

Experimental Results

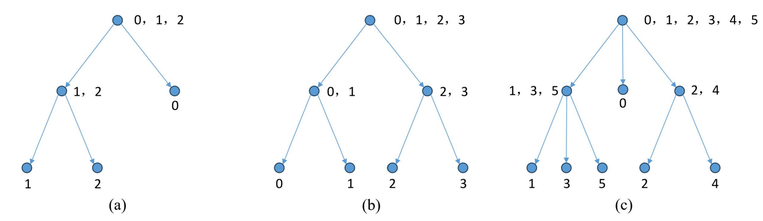
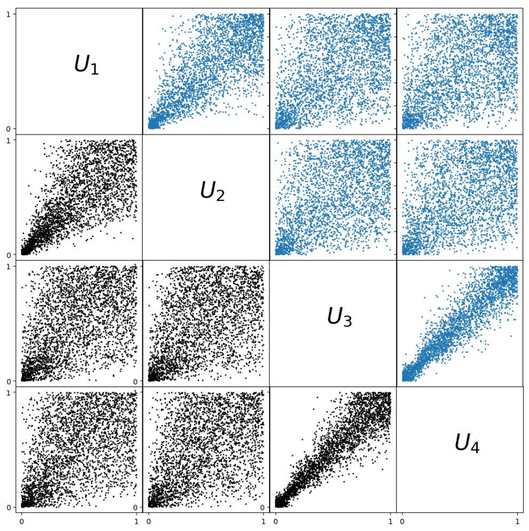
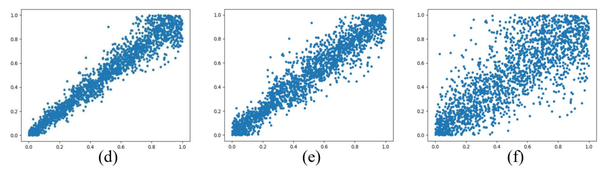


Figure 4: The HAC hierarchy determined using the HACSurv framework: (a) Framingham and SEER datasets, (b) synthetic dataset, and (c) MIMIC-III dataset.





(d): copula between sepsis and ARF
(e): copula between sepsis and Pneumonia
(f): copula between ICH and NSTEMI

Method	Risk 1	Risk 2	Risk 3
DeepHit	0.375 ±0.015	0.382 ±0.015	0.152 ±0.012
DSM	0.243 ±0.008	0.289 ±0.005	0.067 ±0.005
DeSurv	0.362 ±0.013	0.371 ±0.012	0.133 ±0.007
NeuralFG	0.354 ±0.012	0.359 ±0.012	0.157 ±0.004
HACSurv (I)	0.204 ±0.006	0.230 ±0.015	0.067 ±0.001
HACSurv (S)	0.039 ±0.005	0.096 ±0.008	0.013 ±0.000
HACSurv	**0.023** ±0.004	**0.012** ±0.002	**0.008** ±0.002

Metric	Model	Synthetic	Framingham	SEER	MIMIC-III
C^{td}-index	DeepHit	0.611 ± 0.008	0.699 ± 0.019	0.791 ± 0.000	0.765 ± 0.015
	DSM	0.603 ± 0.009	0.734 ± 0.016	0.781 ± 0.006	0.752 ± 0.018
	DeSurv	0.617 ± 0.007	0.674 ± 0.022	0.787 ± 0.005	0.755 ± 0.021
	NeuralFG	0.549 ± 0.019	0.746 ± 0.010	0.779 ± 0.004	0.760 ± 0.026
	HACSurv (I)	0.635 ± 0.005	$\mathbf{0.748} \pm 0.009$	0.797 ± 0.003	0.764 ± 0.018
	HACSurv (S)	0.636 ± 0.006	0.746 ± 0.013	0.797 ± 0.003	0.760 ± 0.016
	HACSurv	$\mathbf{0.643} \pm 0.007$	0.747 ± 0.012	$\mathbf{0.798} \pm 0.002$	$\mathbf{0.768} \pm 0.018$
IBS	DeepHit	0.151 ± 0.003	0.089 ± 0.003	0.074 ± 0.001	0.264 ± 0.083
	DSM	0.086 ± 0.003	0.088 ± 0.004	0.069 ± 0.000	0.226 ± 0.080
	DeSurv	0.143 ± 0.002	0.104 ± 0.004	0.068 ± 0.001	0.268 ± 0.085
	NeuralFG	0.173 ± 0.016	0.091 ± 0.003	0.068 ± 0.001	0.276 ± 0.097
	HACSurv (I)	$\mathbf{0.077} \pm 0.002$	0.092 ± 0.005	$\mathbf{0.067} \pm 0.001$	0.169 ± 0.049
	HACSurv (S)	0.078 ± 0.004	0.091 ± 0.006	$\mathbf{0.067} \pm 0.001$	$\mathbf{0.131} \pm 0.021$
	HACSurv	0.100 ± 0.008	$\mathbf{0.085} \pm 0.003$	$\mathbf{0.067} \pm 0.001$	0.163 ± 0.068

- Synthetic: HACSurv successfully approximates the underlying dependency structure and reduces bias in the estimated marginal survival distributions.
- Real-world: The dependency identified by HACSurv appears to be consistent with common medical knowledge. It achieves superior performance in terms of the C-index.
- Since IBS relies on the independent censoring assumption, it's not a proper scoring rule.

Main References

[1] Nelsen, R. B. **An Introduction to Copulas**, *Springer*, 2nd ed., 2007.

[2] Ng, Y., Hasan, A., Elkhalil, K., & Tarokh, V. **Generative Archimedean Copulas**, *UAI*, 2021.

[3] Zhang, W., Ling, C. K., & Zhang, X. **Deep Copula-Based Survival Analysis for Dependent Censoring with Identifiability Guarantees**, *AAAI*, 2024.