Copula Link-Based Additive Models for Bivariate Time-to-Event Outcomes with General Censoring Scheme

Danilo Petti [1] Alessia Eletti [2] Giampiero Marra [2] Rosalba Radice [3]

[1]University of Salerno [2]University College London [3]Bayes Business School

May 26, 2022

[This paper is under review: Computational Statistics & Data Analysis]

Overview

Motivation

- Difficulty in diagnosing
- Treatments

Flexibility of the proposed model

- Flexibility
- Computational advantage

Data Analysis

- Simulation study
- Case study

Motivation

Two class of motivation:

Age- Related Macula degeneration

- Difficulty diagnosing
- Treatments

Link-based additive survival model

- Flexibility
- Computational advantages

Difficulty diagnosing

- No warning symptoms,making the diagnosis difficult
- In both forms of the AMD, the wet and the dry form, if diagnosed early the **damage** to the ocular system can be **limited**

Risky treatments

- The main treatment against the AMD is based on **steroids** whose have an anti-inflammatory effects

  Potentially **dangerous side effects**!!

  - **Dexamethasone, Triamcinolone**: Can not be used in case of pathologies and in pregnant women. Not to mention the side effects

- **Intravitreal injections**, the last frontier in the fight against AMD. **Have to be repeated every month**!!

The Model

The survival times $S(\cdot)$ linked via a copula $C(\cdot)$.

$$S(t_{1i}, t_{2i}|\mathbf{x}_i; \boldsymbol{\delta}) = C\left(S_1(t_{1i}|\mathbf{x}_{1i}; \boldsymbol{\beta}_1), S_2(t_{2i}|\mathbf{x}_{2i}; \boldsymbol{\beta}_2); m\left\{\eta_{3i}(\mathbf{x}_{3i}; \boldsymbol{\beta}_3)\right\}\right),$$

The marginal survival functions can be written as

$$g_v\left[S(t_{vi}|\mathbf{x}_{vi}; \boldsymbol{\beta})\right] = \eta_{vi}(t_{vi}, \mathbf{x}_{vi}; \boldsymbol{f}_v(\boldsymbol{\beta}_v)),$$

where $g_v : (0,1) \to \mathbb{R}$ is link function, $\eta_{vi}(t_{vi}, \mathbf{x}_{vi}; \boldsymbol{f}_v(\boldsymbol{\beta}_v)) \in \mathbb{R}$ is an additive predictor

$$\eta_{\nu i} = \beta_{\nu 0} + \sum_{k_\nu=1}^{K_\nu} s_{\nu k_\nu}(\mathbf{z}_{\nu k_\nu i}), \; i = 1, \ldots, n, \; \nu = 1, 2, 3$$

Flexibility

The only limit is sample size

- No censoring constraints
- Vast selection of copulas (included rotation) implemented: Clayton ("`C0`"), Frank ("`F`"), Gaussian ("`N`"), Gumbel ("`G0`")
- Proportional hazard `PH`, Proportional Odds `PO`, Probit `probit` link specifiable
- Everything close at R in GJRM package

expr	min	lq	mean	median	uq	max
δ_{UU}	6.30	6.40	6.55	6.48	6.56	7.11
δ_{RR}	6.26	6.37	6.43	6.42	6.47	6.94
δ_{LL}	6.27	6.38	6.45	6.43	6.52	7.01
δ_{II}	6.29	6.38	6.47	6.46	6.52	7.03
δ_{UR}	6.20	6.32	6.41	6.38	6.43	7.28
δ_{RU}	6.21	6.31	6.37	6.36	6.43	6.92
δ_{UL}	6.21	6.32	6.38	6.36	6.43	6.94
δ_{LU}	6.23	6.29	6.36	6.33	6.40	6.95
δ_{UI}	6.20	6.30	6.37	6.36	6.43	6.96
δ_{IU}	6.19	6.31	6.38	6.36	6.42	6.96
δ_{RL}	6.20	6.29	6.38	6.35	6.43	7.25
δ_{LR}	6.19	6.28	6.36	6.34	6.40	6.88
δ_{RI}	6.21	6.30	6.37	6.35	6.41	6.98
δ_{IR}	6.20	6.29	6.45	6.38	6.52	7.36
δ_{LI}	6.35	6.52	6.70	6.62	6.75	8.07
δ_{IL}	6.35	6.73	6.88	6.89	7.03	7.43

Table B.8: Computational time in seconds Analytical Hessian

expr	min	lq	mean	median	uq	max
δ_{UU}	1226.32	1226.52	1255.69	1226.71	1270.37	1314.03
δ_{RR}	1227.28	1230.43	1232.60	1233.58	1235.27	1236.95
δ_{LL}	1234.69	1236.02	1236.56	1237.35	1237.50	1237.64
δ_{II}	1255.28	1256.15	1257.47	1257.03	1258.57	1260.10
δ_{UR}	1232.31	1232.85	1233.89	1233.38	1234.68	1235.99
δ_{RU}	1232.30	1232.64	1233.16	1232.98	1233.59	1234.19
δ_{UL}	1230.06	1231.70	1232.68	1233.34	1233.99	1234.64
δ_{LU}	1234.68	1236.05	1236.59	1237.43	1237.54	1237.66
δ_{UI}	1234.54	1239.75	1241.89	1244.97	1245.56	1246.15
δ_{IU}	1244.39	1244.59	1244.96	1244.78	1245.24	1245.70
δ_{RL}	1243.13	1243.59	1244.34	1244.06	1244.95	1245.84
δ_{LR}	1241.44	1241.66	1242.43	1241.88	1242.92	1243.96
δ_{RI}	1236.53	1238.15	1238.84	1239.78	1240.00	1240.22
δ_{IR}	1235.49	1235.50	1236.16	1235.51	1236.49	1237.48
δ_{LI}	1232.93	1233.78	1234.96	1234.63	1235.98	1237.34
δ_{IL}	1235.25	1236.70	1242.62	1238.15	1246.31	1254.46

Table B.9: Computational time in seconds Numerical Hessian

Computing times (seconds) **analytical** (left) **numerical** (right) hessian.

Model: t-student ("T") copula, proportional hazard ("PH") as the first margin and proportional odds ("PO") as second

Simulation study: setup

Setup Details

$$T_{1i} = \log[-\log S_{10}(t_{1i})] + \beta_{11} z_{1i} + s_{11}(z_{21})$$
$$T_{2i} = \log\left[\frac{\{1 - S_{20}(t_{2i})\}}{S_{20}(t_{2i})}\right] + \beta_{21} z_{1i} + \beta_{22} z_{3i}$$
$$\eta_{3i} = \beta_{31} z_{1i} + s_{31}(z_{2i})$$

Survivals:
$S_{10i}(t_{1i}) = 0.9 \exp\left\{(-0.4 t_{1i}^{2.5})\right\}$,
$S_{20}(t_{2i}) = S_{10}(t_{2i}) = 0.9 \exp\left\{(-0.4 t_{1i}^{2.5})\right\}$
Smooth functions:
$s_{11}(z_i) = \sin(2\pi z_i)$, $s_{31} = 3\sin(\pi z_i)$
Linear parameter:
$\beta_{11} = -1.5$, $\beta_{21} = \beta_{22} = 1.2$, $\beta_{31} = -1.5$.

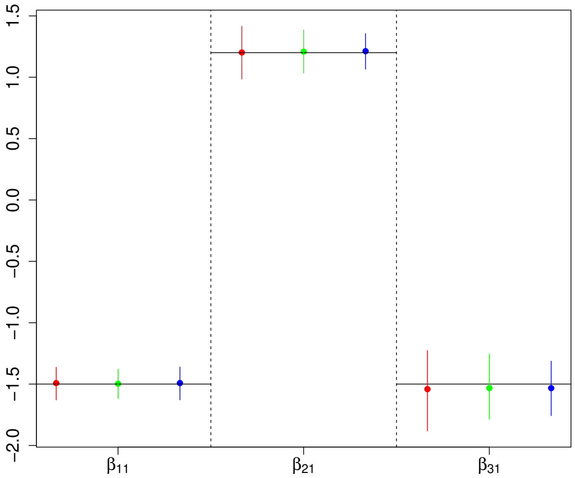

Figure 1: Mild Censoring Results (62.86% and 44.98%): Circles are mean estimate, bars are the estimates' range (5%-95% quantiles). True values are represented by black solid lines. Circles and vertical bar refer to the results obtained for n=1000 , n=1500, n=2000

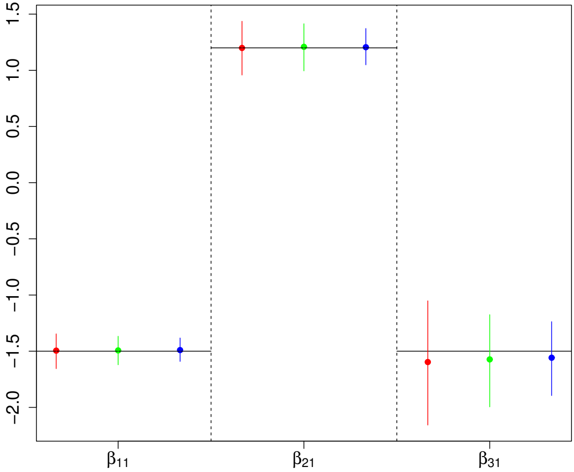

Figure 2: High Censoring Results (84.82% and 77.13%): Circles are mean estimate, bars are the estimates' range (5%-95% quantiles). True values are represented by black solid lines. Circles and vertical bar refer to the results obtained for n=1000 , n=1500, n=2000

AREDS dataset

t11	t12	t21	t22	SNP	Sev1E	Sev2E	Age	Cens
0.00	2.00	0.00	2.00	1	6	8	67	II
0.00	2.00	5.90	9.30	0	7	4	68	II
8.00	9.10	10.00	NA	0	7	7	65	IR
3.00	4.10	3.00	4.10	1	6	6	64	II
4.80	5.80	0.00	1.80	1	7	7	68	II
10.00	NA	10.00	NA	0	6	6	74	RR

Sample of 600 participants between 55-80 years, free of any illness or medical condition. Data affected by Interval and Right censoring.

AREDS data: Estimation Results

Model: PlanketPL Copula and Proportional Hazards PO

	First Eye		Second Eye	
	Estim. (SE)	Pvalue	Estimate (SE)	Pvalue
(Intercept)	-18.0368(4.39)	***	-33.2811 (10.89)	**
Age	-	-	0.0364 (0.01)	**
SevScale5	0.6707 (0.24)	**	0.8187 (0.25)	**
SevScale6	1.0049 (0.22)	***	1.2957 (0.23)	***
SevScale7	1.9255 (0.23)	***	2.4270 (0.25)	***
SevScale8	2.8208 (0.31)	***	3.2793 (0.32)	***
SNB1	0.3269 (0.16)	**	0.4589 (0.16)	**
SNB2	0.6058 (0.23)	**	0.7874 (0.22)	**

Table 1: AREDS data. Parameters estimates and standard error resulted from model fitting using gjrm().

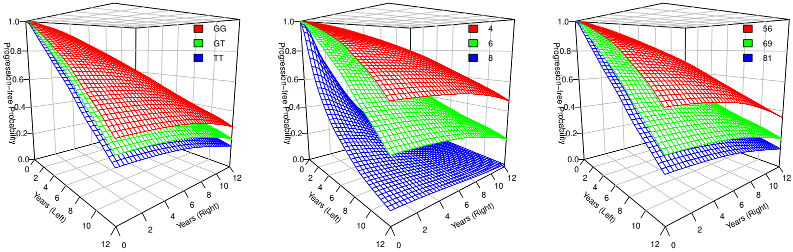

Figure 3: **left panel**, age is set to 69, and AMD severity score to 6 for both eyes. **Middle panel**, age is set to 69, and genotype to GT. **Right panel**, genotype is set to GT, and AMD severity score to 6 in both eyes.

Key Findigs

- Two Eyes does not progress simultaneously. Organs are kinda dependent (Copula)
- Severity Score and Age have a strong impact on AMD
- Genetic Factors impacts less than we expected

Thank you