Individual Claims Reserving with dependent censored data

Marie Michaelides

michaelides.marie@uqam.ca

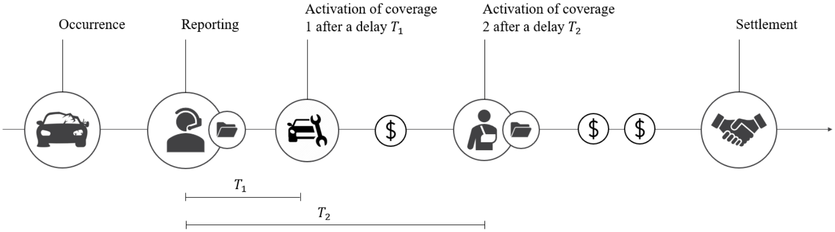
Joint work with Mathieu Pigeon (UQAM) and Hélène Cossette (ULaval)

July 2023

Contents

1. Introduction

- Consider a policy providing two insurance coverages.
- $\boldsymbol{T} = (T_1, T_2)$: activation delays for both coverages.



If a claim activates coverage 1, will it also activate coverage 2? If so, how long after?

$\Longrightarrow$ We work with **censored dependent** data.

2. Some reminders on Archimedean copulas

Archimedean copulas are characterized by a generator $\phi(.) : [0,1] \to [0,\infty]$ with $\phi(1) = 0$:

$$C(u_1, u_2) = \phi^{-1}(\phi(u_1) + \phi(u_2))$$

Given a copula C, we can retrieve Kendall's tau

$$\tau(T_1, T_2) = 4 \int_0^1 \int_0^1 C(u_1, u_2) dC(u_1, u_2) - 1.$$

Table 1: Most commonly used bivariate Archimedean copulas. $\tilde{u} = -\ln u$ and $\bar{u} = 1 - u$.

Copula	$C_\alpha(u_1, u_2)$	$\phi_\alpha(t)$	τ
Clayton	$(u_1^{-\alpha} + u_2^{-\alpha} - 1)^{-1/\alpha}$	$t^{-\alpha} - 1$	$\alpha/(\alpha+2)$
Frank	$-\frac{1}{\alpha} \ln\left(1 + \frac{(e^{-\alpha u_1} - 1)(e^{-\alpha u_2} - 1)}{e^{-\alpha} - 1}\right)$	$-\ln\left(\frac{e^{-\alpha t} - 1}{e^{-\alpha} - 1}\right)$	$1 + \frac{4}{\alpha}\left(\int_0^\alpha \frac{\xi}{\alpha(e^\xi - 1)} d\xi - 1\right)$
Gumbel-Hougaard	$\exp\left(-[(\tilde{u}_1)^\alpha + (\tilde{u}_2)^\alpha]^{1/\alpha}\right)$	$(-\ln t)^\alpha$	$1 - 1/\alpha$
Joe	$1 - (\bar{u}_1^\alpha + \bar{u}_2^\alpha - \bar{u}_1^\alpha \bar{u}_2^\alpha)^{1/\alpha}$	$-\ln(1 - (1-t)^\alpha)$	-

Nonparametric estimation of the generator

Let $Z = C_\phi(U_1, U_2)$. The distribution K of Z is given by

$$K(z) = z - \lambda(z),$$

with

$$\lambda(\nu) = \frac{\phi(\nu)}{\phi^{(1)}(\nu)},$$

for $0 < \nu \leq 1$. The generator can then be recovered by solving

$$\phi(\nu) = \exp\left\{ \int_{\nu_0}^{\nu} \frac{1}{\lambda(t)} dt \right\},$$

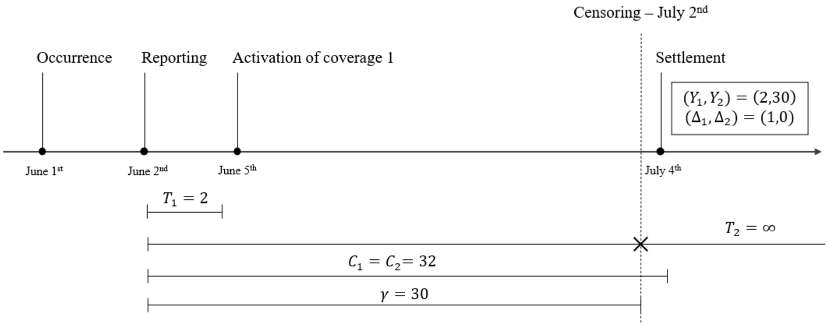
with $0 < \nu_0 < 1$ an arbitrarily chosen constant. In the Archimedean case, Kendall's tau can then be computed using

$$\tau = 4 \int_0^1 \lambda(\nu) d\nu + 1 = 3 - 4 \int_0^1 K(\nu) d\nu.$$

3. Statistical Model

Notation:

- $\boldsymbol{T} = (T_1, T_2)$: vector of true survival times (*activation delays*)
- $\boldsymbol{C} = (C_1, C_2)$ with $C_1 = C_2$: vector of censoring variables (*settlement delays*)
- γ_i for $i = 1, 2$: limits imposed on the survival times
- $\boldsymbol{Y} = (Y_1, Y_2)$ with $Y_i = \min(T_i, C_i, \gamma_i)$ for $i = 1, 2$: observed times
- $\boldsymbol{\Delta} = (1[Y_1 = T_1], 1[Y_2 = T_2])$: censoring indicators



Goal: find the nonparametric estimator $\hat{\phi}(\nu)$ of the copula generator for bivariate censored data:

$$\hat{\phi}(\nu) = \exp\left\{ \int_{\nu_0}^{\nu} \frac{1}{t - \hat{K}(t)} dt \right\}$$

Step 1: Extension of **Beran** (1981)'s estimator

$$\hat{F}_{1|2}(y_1|y_2) = 1 - \prod_{Y_{i1} \le y_1, \Delta_{i1}=1} \left(1 - \frac{W_{ni2}(y_2; h_n)}{\sum_{j=1}^{n} W_{nj2}(y_2; h_n) \, 1_{Y_{j1} \ge Y_{i1}}} \right)$$

Step 2: Joint distribution $\hat{F}(\boldsymbol{y})$

$$\hat{F}(\boldsymbol{y}) = w(\boldsymbol{y}) \int_0^{y_2} \hat{F}_{1|2}(y_1|z_2) d\tilde{F}_2(z_2) + (1 - w(\boldsymbol{y})) \int_0^{y_1} \hat{F}_{2|1}(y_2|z_1) d\tilde{F}_1(z_1)$$

with $\tilde{F}_j(z_j)$ the marginal estimators of **Kaplan-Meier** (1958).

Step 3: Wang and Wells (2000)'s estimator for $\hat{K}(t)$

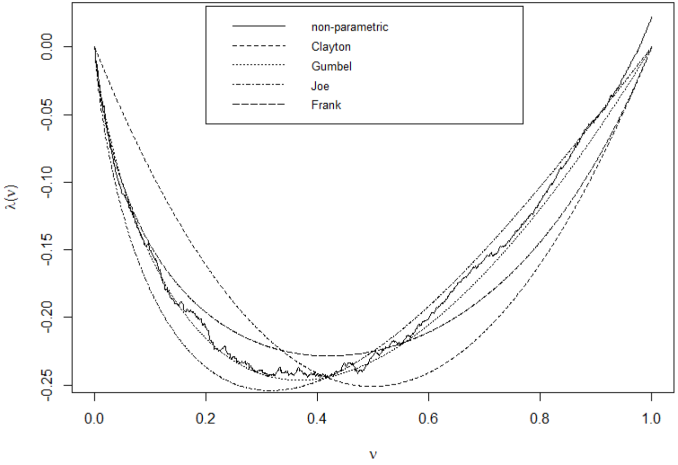
$$\hat{K}(t) = \int_0^\infty \int_0^\infty 1[\hat{F}(\boldsymbol{y}) \le t] d\hat{F}(\boldsymbol{y})$$

Step 4: Comparison with selected copula models using $\hat{\lambda}(.)$

Graphical comparison

Knowing $\hat{K}(\nu)$, estimate Kendall's tau and retrieve the copula parameters $\hat{\alpha} = g^{-1}(\hat{\tau})$.

Compare the plot of $\hat{\lambda}(\nu) = \nu - \hat{K}(\nu)$ to those of $\hat{\lambda}_{\hat{\alpha}}(\nu)$.



4. Model validation

UQÀM | co-operators

Chaire Co-operators en analyse des risques actuariels

Omnibus procedure

- Likelihood function:

$$L(u_1, u_2, \delta_1, \delta_2; \alpha) = \prod_{i=1}^{n} c(u_{1i}, u_{2i}; \alpha)^{\delta_{1i}\delta_{2i}} + \left(\frac{\partial C(u_{1i}, u_{2i}; \alpha)}{\partial u_1}\right)^{\delta_{1i}(1-\delta_{2i})} + \left(\frac{\partial C(u_{1i}, u_{2i}; \alpha)}{\partial u_2}\right)^{(1-\delta_{1i})\delta_{2i}} + C(u_{1i}, u_{2i}; \alpha)^{(1-\delta_{1i})(1-\delta_{2i})}.$$

- Optimal dependence parameter value to compare to that found using $\hat{K}(\nu)$:

$$\hat{\alpha}^* = \arg\max L(u_1, u_2, \delta_1, \delta_2; \alpha)$$

L^2-norm

$$S(\hat{\alpha}) = \int_0^1 \left(\hat{K}(\nu) - K_{\hat{\alpha}}(\nu)\right)^2 dK_{\hat{\alpha}}(\nu).$$

Riemann sum approximate:

$$\hat{S}(\hat{\alpha}) = \sum_{i=1}^{n} \left(\hat{K}(\nu_{(i)}) - K_{\hat{\alpha}}(\nu_{(i)})\right)^2 (\nu_{(i)} - \nu_{(i-1)}).$$

Copula 1 provides the closest fit to the data among M copulas under H_0:

$$H_0 : \min_{k=2,\ldots,M} \hat{S}(\hat{\alpha}_k) - \hat{S}(\hat{\alpha}_1) > 0$$

$$H_1 : \min_{k=2,\ldots,M} \hat{S}(\hat{\alpha}_k) - \hat{S}(\hat{\alpha}_1) \leq 0$$

Wang (2010)'s goodness-of-fit test

Let $C_\alpha(u,v) = \phi_\alpha^{-1}\left[\phi_\alpha(u) + \phi_\alpha(v)\right]$ be an Archimedean copula. Then,

$$U = \frac{\phi(u)}{\phi[C(u,v)]}, \quad V = C(u,v)$$

are shown to be independent by Genest and Rivest (1993), leading to the test

$$H_0 : \rho = 0 \quad \text{vs} \quad H_1 : \rho \neq 0.$$

Let

$$r_n = \frac{\sum_{i=1}^n (\hat{U}_i - \bar{\hat{U}})(\hat{V}_i - \bar{\hat{V}})}{\sqrt{\sum_{i=1}^n (\hat{U}_i - \bar{\hat{U}})^2 \sum_{i=1}^n (\hat{V}_i - \bar{\hat{V}})^2}}$$

The test statistic is defined as

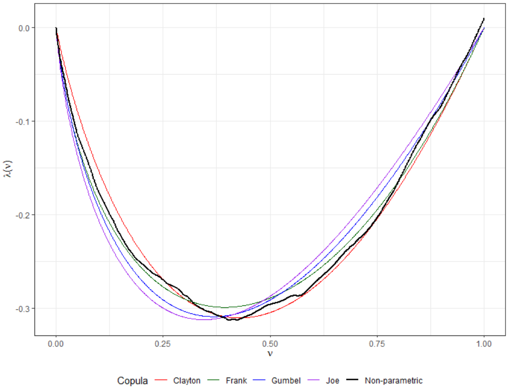
$$Z_n = \frac{1}{2} \log\left[\frac{1 + r_n}{1 - r_n}\right]$$

and we have that $\sqrt{n} Z_n \to N(0,1)$ in distribution. This test can be extended to censored data by adjusting the distributions of $\hat{U}$ and $\hat{V}$.

5. Simulation Study

Setup:

- $n = 500$ observations
- Clayton copula with $\tau = 0.25$
- Exponential(1) marginals
- $\pm 20\%$ censoring



- **Omnibus procedure**: comparison of $\hat{\alpha}$ with $\hat{\alpha}^* = \arg\max L(u_1, u_2, \delta_1, \delta_2; \alpha)$

Copula	$\hat{\alpha}$	$\hat{\alpha}^*$
Clayton	0.723	0.767
Frank	2.540	2.156
Gumbel	1.361	1.190
Joe	1.650	1.127

- **Wang (2010)**'s goodness-of-fit test for censored data

Table 1: Percentage of rejection of the null hypothesis for different copulas.

True copula	τ	Copula under H_0 Clayton p_W	Gumbel p_W
Clayton	0.2	0.10	0.64
	0.4	0.20	0.94
Gumbel	0.2	0.90	0.14
	0.4	0.72	0.25

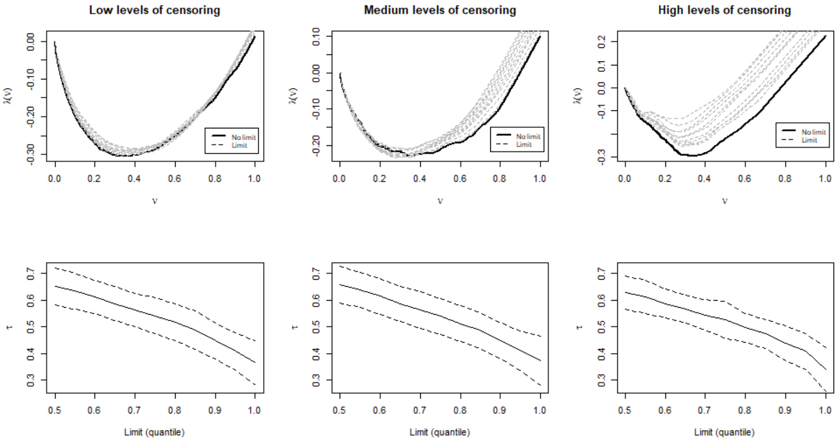
- L^2**-norm** test

Table 2: Percentage of rejection of the null hypothesis for different copulas

True copula	τ	Copula under H_0 Clayton	Frank	Gumbel	Joe
		$ps(\hat{\alpha})$	$ps(\hat{\alpha})$	$ps(\hat{\alpha})$	$ps(\hat{\alpha})$
Frank	0.2	1.00	0.20	0.80	1.00
	0.4	1.00	0.36	0.64	1.00
	0.6	1.00	0.34	0.70	0.96
Gumbel	0.2	1.00	1.00	0.15	0.85
	0.4	1.00	0.96	0.14	0.90
	0.6	1.00	0.98	0.02	1.00
Joe	0.2	1.00	1.00	0.82	0.18
	0.4	1.00	1.00	0.84	0.16
	0.6	1.00	1.00	0.84	0.16

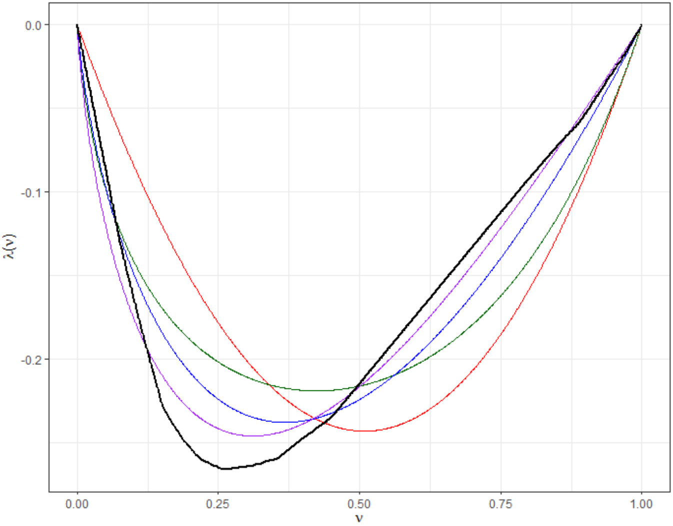
Impact of decreasing γ

Sample of $n = 500$ observations from a Clayton Copula with $\alpha = 0.7$ and margins such that $X_1 \sim \text{LogNormal}(\mu_1 = 8, \sigma_1 = 1)$ and $X_2 \sim \text{LogNormal}(\mu_2 = 7, \sigma_2 = 3)$.



6. Application to claims reserving

- Automobile insurance dataset with 2 insurance coverages
- Limit: $\gamma_1 = \gamma_2 = 730$ days



Thank you!

Selected references

Akritas, M. (1994), Nearest neighbor estimation of a bivariate distribution under random censoring, *Annals of Statistics*, 22, 1299-1327.

Akritas, M.G., Van Keilegom, I. (2003), Estimation of Bivariate and Marginal Distributions with Censored Data, *Journal of the Royal Statistical Society Series B: Statistical Methodology*, 65:2, 457-471.

Beran, R. (1981), Nonparametric Regression with Randomly Censored Survival Data, *Berkeley, CA: University of California*.

Denuit, M., Purcaru, O., Van Keilegom, I. (2006), Bivariate Archimedean copula modelling for loss-ALAE data in non-life insurance, *Journal of Actuarial Practice*, 5, 1993-2006.

Genest, C., Rivest, L. (1993), Statistical Inference Procedures for Bivariate Archimedean Copulas, *Journal of the American Statistical Association*, 88, 1034-1043.

Wang, W., Wells, M. T. (2000), Model selection and semi-parametric inference for bivariate censored data, *Journal of the American Statistical Association*, 95, 62-72.