Parameterization of copula functions for bivariate survival data in the \texttt{surrosurv} package (v. 1.1.25).

Motelling and simulation

Federico Rotolo

January 18, 2019

Let define the joint survival function of $S$ and $T$ via a copula function:

$$S(s,t) = P(S > s, T > t) = C(u,v)|_{u=S_S(s), v=S_T(t)}, \tag{1}$$

where $S_S(\cdot) = P(S > s)$ and $S_T(\cdot) = P(T > t)$ are the marginal survival functions of $S$ and $T$.

**Modelling**

In the case of possibly right-censored data, the individual contribution to the likelihood is

- $S(s,t) = C(u,v)|_{S_S(s), S_T(t)}$ if $S$ is censored at time $s$ and $T$ is censored at time $t$,
- $-\frac{\partial}{\partial t} S(s,t) = \frac{\partial}{\partial v} C(u,v)|_{S_S(s), S_T(t)} f_T(t)$ if $S$ is censored at time $s$ and $T = t$,
- $-\frac{\partial}{\partial s} S(s,t) = \frac{\partial}{\partial u} C(u,v)|_{S_S(s), S_T(t)} f_S(s)$ if $S = s$ and $T$ is censored at time $t$,
- $\frac{\partial^2}{\partial s \partial t} S(s,t) = \frac{\partial^2}{\partial u \partial v} C(u,v)|_{S_S(s), S_T(t)} f_S(s)f_T(t)$ if $S = s$ and $T = t$.

**Clayton copula**

The bivariate copula is defined as

$$C(u,v) = (u^{-\theta} + v^{-\theta} - 1)^{-1/\theta}, \quad \theta > 0. \tag{2}$$

The first derivative with respect to $u$ is

$$\frac{\partial}{\partial u} C(u,v) = (u^{-\theta} + v^{-\theta} - 1)^{-\frac{1+\theta}{\theta}} u^{-(1+\theta)} \left[ \frac{C(u,v)}{u} \right]^{1+\theta}. \tag{3}$$

The second derivative with respect to $u$ and $v$ is

$$\frac{\partial^2}{\partial u \partial v} C(u,v) = (1 + \theta) \frac{C(u,v)^{1+2\theta}}{(uv)^{1+\theta}}. \tag{4}$$

The \textit{\theta}'s tau for the Clayton copula is

$$\tau = \frac{\theta}{\theta + 2}. \tag{5}$$

**Plackett copula**

The bivariate copula is defined as

$$C(u,v) = \left[ \frac{Q - R^{1/2}}{2(\theta - 1)} \right], \quad \theta > 0, \tag{6}$$
with
\[ Q = 1 + (\theta - 1)(u + v), \]
\[ R = Q^2 - 4\theta(\theta - 1)uv. \]  \hspace{1cm} (7)

Given that
\[ \frac{\partial}{\partial u} Q = \theta - 1, \] \hspace{1cm} (8)
\[ \frac{\partial}{\partial u} R = 2(\theta - 1)(1 - (\theta + 1)v + (\theta - 1)u) \]
\[ = 2(\theta - 1)(Q - 2\theta v), \] \hspace{1cm} (9)

the first derivative of \( C(u, v) \) with respect to \( u \) is
\[ \frac{\partial}{\partial u} C(u, v) = \frac{1}{2} \left[ 1 - \frac{1 - (\theta + 1)v + (\theta - 1)u}{R^{1/2}} \right] \]
\[ = \frac{1}{2} \left[ 1 - \frac{Q - 2\theta v}{R^{1/2}} \right]. \] \hspace{1cm} (10)

By defining
\[ f = Q - 2\theta v, \] \hspace{1cm} (11)
\[ g = R^{1/2} \] \hspace{1cm} (12)
and given that
\[ f' = \frac{\partial}{\partial v} f = -\theta + 1, \] \hspace{1cm} (13)
\[ g' = \frac{\partial}{\partial v} g = \frac{\theta - 1}{R^{1/2}} \left( 1 - (\theta + 1)u + (\theta - 1)v \right) \]
\[ = \frac{\theta - 1}{R^{1/2}} \left( Q - 2\theta u \right), \] \hspace{1cm} (14)
then, the second derivative with respect to \( u \) and \( v \) is (see Appendix A for full details)
\[ \frac{\partial^2}{\partial u \partial v} C(u, v) = -\frac{f'g - fg'}{2g^2} \]
\[ = \frac{\theta}{R^{3/2}} \left[ 1 + (\theta - 1)(u + v - 2uv) \right] \]
\[ = \frac{\theta}{R^{3/2}} \left[ Q - 2(\theta - 1)uv \right]. \] \hspace{1cm} (15)

The Kendall’s tau for the Plackett copula cannot be computed analytically and is obtained numerically.

**Gumbel–Hougaard copula**

The bivariate copula is defined as
\[ C(u, v) = \exp \left( -Q^\theta \right), \quad \theta \in (0, 1), \] \hspace{1cm} (16)
with
\[ Q = (-\ln u)^{1/\theta} + (-\ln v)^{1/\theta}. \] \hspace{1cm} (17)

Given that
\[ \frac{\partial}{\partial u} Q = -\frac{(-\ln u)^{1/\theta - 1}}{\theta u}, \] \hspace{1cm} (18)
then, the first derivative with respect to \( u \) is
\[ \frac{\partial}{\partial u} C(u, v) = \frac{(-\ln u)^{1/\theta - 1}}{u} C(u, v) Q^{\theta - 1} \] \hspace{1cm} (19)
and the second derivative with respect to \( u \) and \( v \) is
\[ \frac{\partial^2}{\partial u \partial v} C(u, v) = \left[ (-\ln u)(-\ln v) \right]^{1/\theta - 1} \]
\[ \frac{(-\ln u)(-\ln v)}{uv} C(u, v) Q^{\theta - 2} \left[ \frac{1}{\theta} - 1 + Q^\theta \right]. \] \hspace{1cm} (20)

The Kendall’s tau for the Gumbel–Hougaard copula is
\[ \tau = 1 - \theta. \] \hspace{1cm} (21)
Simulation

Clayton copula

The function `simData.ccc()` generates data from a Clayton copula model. First, the time value for the surrogate endpoint $S$ is generated from its (exponential) marginal survival function:

$$ S = -\log(U_S/\lambda_S), \quad \text{with } U_S := S_S(S) \sim U(0,1). \quad (22) $$

Then, the time value for the true endpoint $T$ is generated conditionally on the value $s$ of $S$. The conditional survival function of $T \mid S$ is

$$ S_{T|S}(t \mid s) = \frac{\partial}{\partial s} S(s, t) = \frac{\partial}{\partial u} C(u, v) - \frac{\partial}{\partial s} S(s, 0) = \frac{\partial}{\partial u} C(u, 1) \quad (23) $$

As the Clayton copula is used, we get (see Equation 3)

$$ S_{T|S}(t \mid s) = \left[ \frac{C(S_S(s), S_T(t))}{C(S_S(s), 1)} \right]^{1+\theta} = \left[ \frac{U_S^{-\theta} + S_T(t)^{-\theta} - 1}{U_S^{-\theta}} \right]^{-\frac{1+\theta}{\theta}} $$

$$ = \left[ 1 + U_S^\theta (S_T(t)^{-\theta} - 1) \right]^{-\frac{1+\theta}{\theta}} \quad (24) $$

By generating uniform random values for $U_T := S_{T|S}(T \mid s) \sim U(0,1)$, the values for $T \mid S$ are obtained as follows:

$$ U_T = \left[ 1 + U_S^\theta (S_T(T)^{-\theta} - 1) \right]^{-\frac{1+\theta}{\theta}} $$

$$ S_T(T) = \left[ \left( U_T^{-\frac{\theta}{1-\theta}} - 1 \right) U_S^{-\theta} + 1 \right]^{-1/\theta} $$

$$ T = -\log(S_T(T)/\lambda_T). \quad (25) $$

Gumbel–Hougaard copula

The function `simData.gh()` generates data from a Gumbel-Hougaard copula model. First, the time value for the surrogate endpoint $S$ is generated from its (exponential) marginal survival function:

$$ S = -\log(U_S/\lambda_S), \quad \text{with } U_S := S_S(S) \sim U(0,1). \quad (26) $$

The conditional survival function of $T \mid S$ is (see Equation 19)

$$ S_{T|S}(t \mid s) = \exp \left( Q(S_S(s), 1)^{\theta} - Q(S_S(s), S_T(t))^{\theta} \right) \left[ \frac{Q(S_S(s), S_T(t))^{\theta}}{Q(S_S(s), 1)} \right]^{\theta-1} $$

$$ = \exp \left( - \log U_S - \left[ (-\log U_S)^{\frac{1}{\theta}} + (-\log S_T(T))^{\frac{1}{\theta}} \right]^{\theta} \left[ 1 + \left( \frac{\log S_T(T)}{\log U_S} \right)^{\frac{1}{\theta}} \right]^{\theta-1} \right) \quad (27) $$

By generating uniform random values for $U_T := S_{T|S}(T \mid s) \sim U(0,1)$, the values for $S_T(T)$ are obtained by numerically solving

$$ U_T - \exp \left( - \log U_S - \left[ (-\log U_S)^{\frac{1}{\theta}} + (-\log S_T(T))^{\frac{1}{\theta}} \right]^{\theta} \left[ 1 + \left( \frac{\log S_T(T)}{\log U_S} \right)^{\frac{1}{\theta}} \right]^{\theta-1} \right) = 0 \quad (28) $$

and then the times $T \mid S$ are

$$ T = -\log(S_T(T)/\lambda_T). \quad (29) $$
A Second Derivative of the Plackett Copula

Let $f = Q - 2\theta v$, and $g = R^{1/2}$, with $Q = 1 + (\theta - 1)(u + v)$ and $R = Q^2 - 4\theta (\theta - 1)uv$. Hence,

\[ f' = \frac{\partial}{\partial v} f = -(\theta + 1), \quad (30) \]
\[ g' = \frac{\partial}{\partial v} g = \frac{\theta - 1}{R^{1/2}} (Q - 2\theta u). \quad (31) \]

Then, the second derivative of $C(u, v)$ with respect to $u$ and $v$ is

\[
\frac{\partial^2}{\partial u \partial v} C(u, v) = -\frac{f'g - fg'}{2g^2} = \frac{fg' - f'g}{2g^2} \\
= \frac{1}{R} \left[ \frac{\theta - 1}{2R^{1/2}} (Q - 2\theta u)(Q - 2\theta v) + (\theta + 1)R^{1/2} \right] \\
= \frac{1}{2R^{3/2}} \left[ (\theta - 1)(Q - 2\theta u)(Q - 2\theta v) + (\theta + 1)R \right] \\
= \frac{1}{2R^{3/2}} \left[ (\theta - 1)(Q^2 + 4\theta^2 uv - 2\theta Q(u + v)) + (\theta + 1)(Q^2 - 4\theta (\theta - 1)uv) \right] \\
= \frac{1}{2R^{3/2}} \left[ (\theta - 1)Q^2 - 4\theta^2(\theta - 1)uv - 2\theta Q(\theta - 1)(u + v) + ((\theta + 1)Q^2 - 4\theta (\theta - 1)uv) \right] \\
= \frac{1}{2R^{3/2}} \left[ (\theta - 1)Q^2 - 4\theta^2(\theta - 1)uv - 2\theta Q(\theta - 1)(u + v) + ((\theta + 1)Q^2 - 4\theta (\theta - 1)uv) \right] (32)
\]

Since $(u + v)(\theta - 1) = Q - 1$, then

\[
\frac{\partial^2}{\partial u \partial v} C(u, v) = \frac{1}{2R^{3/2}} \left[ 2\theta Q^2 - 4\theta(\theta - 1)uv - 2\theta Q(Q - 1) \right] \\
= \frac{1}{2R^{3/2}} \left[ 2\theta Q - 4\theta(\theta - 1)uv \right] \\
= \frac{\theta}{R^{3/2}} \left[ Q - 2(\theta - 1)uv \right]. \quad (33)
\]