‘biSurv’ - an R-package for bivariate survival data

Jeppe Ekstrand Halkjær Madsen

2019-02-28

Testing for independence

Estimating tail dependence for censored data

τ	0	0.25	0.5	0.75
True value	0.100	0.419	0.709	0.891
clayton	0.103	0.421	0.708	0.890
claytonFast	0.147	0.424	0.714	0.894
True value	0.100	0.203	0.385	0.647
Gumbel	0.104	0.213	0.391	0.651
GumbelFast	0.145	0.232	0.387	0.645

Table 1: Lower tail dependence

τ	0	0.25	0.5	0.75
True value	0.200	0.294	0.430	0.648
clayton	0.185	0.273	0.398	0.607
claytonFast	0.147	0.221	0.334	0.535
True value	0.200	0.435	0.647	0.835
Gumbel	0.184	0.411	0.615	0.790
GumbelFast	0.150	0.346	0.549	0.728

Table 2: Upper tail dependence

Estimation of Kendall’s τ for censored data

The function `tauCens` can estimate Kendall’s τ for censored data. This relies on some extrapolation which here is done in a very simple way using the Kaplan-Meier estimator. This means that the extrapolation doesn’t use information from the other failure time when extrapolating, which makes the estimator biased towards 0. The standard error estimate uses an assumption of independence and only takes the variance of the nominator into account, which means that the estimate of the standard error is going to be too small on average. These problems are illustrated with the following simulation. 100 pairs of failure times are simulated from the Clayton copula and the Gumbel copula respectively. The marginal distributions are exponential with a scale parameter of 1, while censoring times are exponentially distributed with a scale parameter of 2 leading to roughly 30 % censoring. Then Kendall’s $\hat{\tau}$ and its standard error are estimated. The simulation is repeated 10000 times and is done for four different values of Kendall’s τ: 0, 0.25, 0.5, 0.75. The results are summarised in Table 2.

The estimator looks unbiased in the case of $\tau = 0$ (which corresponds to the independence copula so the results should be the same for both copulas in this case, which they are), but otherwise biased as expected. The bias seems to be worse for the Gumbel copula than for the Clayton copula. One theory could be that this is because the Gumbel copula has upper tail dependence while the Clayton copula has lower tail dependence so the dependence gets “censored away” in the case of the Gumbel copula, but not in the case of the Clayton copula. We also see that the empirical standard error on average is greater than the estimated one. The estimated standard error is unaffected by the value of Kendall’s τ while the actual standard error is bigger the more Kendall’s τ differs from 0.

τ	0	0.25	0.5	0.75
Clayton	0.007	0.173	0.322	0.440
Clayton theoretical SE	0.035	0.035	0.035	0.035
Clayton empirical SE	0.054	0.070	0.095	0.120
Gumbel	0.007	0.134	0.272	0.407
Gumbel theoretical SE	0.035	0.035	0.035	0.035
Gumbel empirical SE	0.054	0.065	0.085	0.107

Table 3: HUSK AT SKRIVE NOGET HER JEPPE

Conclusion

The estimator is far from perfect, but can still be useful as a quick way of getting an impression of how much (if any) dependence there is in the data.