View metadata, citation and similar papers at core.ac.uk
brought to you by CORE
provided by Research Papers in Economics

CONDITIONS FOR THE ASYMPTOTIC SEMIPARAMETRIC EFFICIENCY OF AN OMNIBUS ESTIMATOR OF DEPENDENCE PARAMETERS IN COPULA MODELS

Christian Genest
Département de mathématiques et de statistique, Université Laval

Bas J. M. Werker
Finance Group, Tilburg University

Abstract Oakes (1994) described in broad terms an omnibus semiparametric procedure for estimating the dependence parameter in a copula model when marginal distributions are treated as (infinite-dimensional) nuisance parameters. The resulting estimator was subsequently shown to be consistent and normally distributed asymptotically (Genest et al. 1995, Shih and Louis 1995). Conditions under which it is also semiparametrically efficient in large samples are given. While these requirements are met for the normal copula model (Klaassen and Wellner 1997), it is argued that this is an exception rather than the norm.

Keywords: Copula model, Efficient influence function, Efficient score function, Semiparametric estimation, Sturm-Liouville differential equations.

AMS subject classification: 62G05, 62G20, 62H99.

1. Introduction

A random vector $(X_1,\ldots,X_m)$ is said to arise from a copula model when its joint distribution function is assumed to be of the form

$$H(x_1,\ldots,x_m) = C_\theta\{F_1(x_1),\ldots,F_m(x_m)\}, \tag{1.1}$$

where $F_i(x_i)$ denotes the marginal distribution of X_i and for each parameter value θ, C_θ is an m-variate distribution function with uniform marginals, i.e., a copula. The same terminology is sometimes used to describe situations in which the joint survival function of the X_i's can be written as

$$\bar{H}(x_1,\ldots,x_m) = C_\theta\{\bar{F}_1(x_1),\ldots,\bar{F}_m(x_m)\} \tag{1.2}$$

in terms of the marginal survival functions $\bar{F}_i(t) = 1 - F_i(t)$.

Representations (1.1) and (1.2) are advantageous in that they allow to model separately the behaviour of each of the X_i's and their dependence structure. Oakes (1989), Bandeen-Roche and Liang (1996), and Klugman and Parsa (1999), among others, have shown the value of this modelling approach in medical, engineering, and actuarial contexts.

In typical applications of copula models, specific parametric forms are selected for the dependence structure (i.e., C_θ) and the marginals (e.g., normal distributions or Pareto survival functions). Then arises the question of estimating the various (scalar or vector-valued) parameters of the model from a random sample $(X_{11},\ldots,X_{m1})$, ..., $(X_{1n},\ldots,X_{mn})$ of data from H or $\bar{H}$, as the case may be. A standard approach consists in maximizing the joint likelihood, either directly or in two steps as suggested in Chapter 10 of Joe (1997). However, inappropriate choices for the marginals could then affect appreciably the estimation of the dependence parameter θ. To avoid this problem, margin-free estimates of θ had to be found. Thus, from ad hoc proposals made in special contexts (e.g., Clayton and Cuzik 1985; Oakes 1982, 1986; Genest 1987; Hougaard 1989) gradually emerged a general strategy for estimating semiparametrically the dependence parameters in copula models.

This omnibus procedure, described in broad terms by Oakes (1994), is quite simple. For each $1 \le i \le m$, let

$$F_{in}(x) = \frac{1}{n+1}\sum_{j=1}^{n} 1\,(X_{ij} \le x)$$

be the (rescaled) empirical distribution function corresponding to X_i, and select the value $\hat{\theta}$ that maximizes the pseudo-log-likelihood

$$\sum_{j=1}^{n} \log c_\theta\left\{F_{1n}(X_{1j}),\ldots,F_{mn}(X_{mj})\right\}, \tag{1.3}$$

where c_θ is the density associated with C_θ, assuming it is strictly positive on $(0,1)^m$. In other words, replace F_i by F_{in} in the likelihood for θ and maximize it. As shown by Genest et al. (1995) and Shih and Louis (1995), the resulting estimator is consistent and asymptotically normally distributed, even in the presence of censorship. This allows for the construction of confidence intervals that have the desired coverage probability asymptotically.

Since this omnibus method is close in spirit to maximum likelihood, it might be hoped that the estimator it yields becomes semiparametrically efficient as the sample size increases. Copula models for which this is indeed the case are characterized in Section 2, where it is argued on intuitive grounds that efficiency should actually be the exception rather than the norm. And indeed, only two instances of semiparametric efficiency are identified in Section 3: the case of independence and the normal copula model, for which this finding was already announced by Klaassen and Wellner (1997). While this list is by no means exhaustive, the search for a more efficient rank-based estimation procedure seems justified, if only on theoretical grounds, and alternative avenues to that effect are briefly touched upon in Section 4. The proof of the characterization result is given in an appendix, using theory described by Bickel et al. (1993).

2. When is the omnibus estimator asymptotically semiparametrically efficient?

Though the arguments developed in this paper are valid in a multivariate context, discussion is restricted hereafter to the case $m = 2$ for ease of exposition. Under sufficient regularity conditions (to be given later), the omnibus semiparametric estimator $\hat{\theta}$ is then a solution of

$$\sum_{j=1}^{n} \dot{\ell}_\theta \left\{F_{1n}(X_{1j}), F_{2n}(X_{2j})\right\} = 0,$$

where $\dot{\ell}_\theta(u_1,u_2) = \partial \log c_\theta(u_1,u_2)/\partial\theta$. While the substitution of consistent estimators F_{in} of the unknown marginals ensures that $\hat{\theta}$ is both rank-based and consistent, this procedure ignores information about the F_i's that may be present in the data when the underlying copula is known to belong to a given class (C_θ).

This observation is not new; e.g., it is at the root of work by Zheng and Klein (1995), who use knowledge about the dependence structure of two competing risks to improve on standard (Kaplan-Meier) estimates of the marginal survival functions. To illustrate this point in a simple way, suppose that the class (C_θ) reduces to a single copula C that concentrates all its probability mass uniformly on the line segments $u_2 = u_1 - 1/2$ and $u_2 = u_1 + 1/2$, according as $u_1 - 1/2 \in [-1/2, 1/2]$ is positive or negative. If random observations $(X_{11}, X_{21}), \ldots, (X_{1n}, X_{2n})$ are drawn from

$$H(x_1,x_2) = C\left\{F_1(x_1), F_2(x_2)\right\}, \tag{2.1}$$

the fact that the "shuffle of min" C is singular (Mikusiński et al. 1992) implies that for all j, one will have either

$$X_{2j} = F_2^{-1}\{F_1(X_{1j}) - 1/2\} \quad \text{or } X_{2j} = F_2^{-1}\{F_1(X_{1j}) + 1/2\},$$

depending on whether X_{1j} is larger or smaller than $\text{med}(X_1) = F_1^{-1}(1/2)$. To estimate the latter, $F_{1n}^{-1}(1/2)$ would then be an inefficient choice. For, the

median of X_1 would necessarily be located in the interval $(X_{1(j^\star)}, X_{1(j^\star+1)})$, where the $X_{1(j)}$'s denote the order statistics of the X_1 sample and $X_{1(j^\star)} = X_{1j}$ with $\text{rank}(X_{2j}) = n$. A fortiori, F_1 should not be estimated by F_{1n}.
If the F_{in}'s are indeed inefficient estimators of the marginals when information is available about the dependence structure, it may then be suspected that the omnibus estimator $\hat{\theta}$ is not always semiparametrically efficient. To determine exactly when it is, results from Chapter 3 of Bickel et al. (1993) can be used. For $i = 1, 2$, define

$$\dot{\ell}_i(u_1, u_2) = \frac{\partial}{\partial u_i} \log c_\theta(u_1, u_2),$$

$$I(\theta) = \int_0^1 \int_0^1 \dot{\ell}_\theta^2(u_1, u_2) c_\theta(u_1, u_2) du_1 du_2,$$

and among the primitives of

$$I_{\theta i}(u_i) = \int_0^1 \dot{\ell}_\theta(u_1, u_2) \dot{\ell}_i(u_1, u_2) c_\theta(u_1, u_2) du_{3-i},$$

let $W_i(u_i)$ denote that one which has zero expectation. A proof of the following proposition may be found in the appendix under the regularity conditions used in Section 3.4 of Bickel et al. (1993).

Proposition. *The omnibus semiparametric estimator* $\hat{\theta}$ *is asymptotically efficient if and only if for each possible value of* θ*, there exists a constant* $0 < I^*(\theta) \leq I(\theta)$ *such that for all* $0 \leq u_1, u_2 \leq 1$,

$$\{1 - I(\theta)/I^*(\theta)\}\, \dot{\ell}_\theta(u_1, u_2) = \sum_{i=1}^{2} \mathrm{E}\{W_i(U_i) \mid U_{3-i} = u_{3-i}\}$$
$$+ \sum_{i=1}^{2} \dot{\ell}_i(u_1, u_2) \int_0^{u_{3-i}} [W_i(v) + \mathrm{E}\{W_{3-i}(U_{3-i}) \mid U_i = v\}]\, dv. \qquad (2.2)$$

Under these conditions, $I^*(\theta)$ *is then the information for estimating* θ *in the semiparametric model.*

3. Illustrations

It is immediate from (2.1) that the estimator $\hat{\theta}$ is semiparametrically efficient and even adaptative at independence. For, in this case, the $\dot{\ell}_i$'s and the $I_{\theta i}$'s vanish, so that the condition of the proposition is trivially verified with $I^*(\theta) = I(\theta)$. Under independence, the empirical distributions are well known to be efficient estimators of the marginals, of course.
The following examples illustrate the calculations involved in checking condition (2.1) for two familiar systems of copulas that include independence as a

special case. The omnibus estimator $\hat{\theta}$ is seen to be semiparametrically efficient in the first model, but not in the second.

Example 1. (Normal distributions). If Φ_θ represents the distribution function of a bivariate normal vector with standardized marginals Φ and correlation coefficient θ, the underlying normal copula is defined on the unit square by

$$C_\theta(u_1,u_2) = \Phi_\theta\left\{\Phi^{-1}(u_1),\Phi^{-1}(u_2)\right\} = \Phi_\theta(v_1,v_2)$$

with $v_i = \Phi^{-1}(u_i)$, $i = 1,2$. Let ϕ denote the standard normal density. Klaassen and Wellner (1997) showed

$$\dot{\ell}_\theta(u_1,u_2) = \frac{v_1 v_2}{1-\theta^2} + \frac{\theta}{1-\theta^2}\left(1 - \frac{v_1^2+v_2^2-2\theta v_1 v_2}{1-\theta^2}\right),$$

$$\dot{\ell}_i(u_i) = \frac{\theta}{1-\theta^2}\,\frac{v_{3-i}-\theta v_i}{\phi(v_i)}, \quad I_{\theta i}(u_i) = \frac{\theta}{1-\theta^2}\,\frac{v_i}{\phi(v_i)},$$

and

$$W_i(u_i) = \frac{1}{2}\,\frac{\theta}{1-\theta^2}\,\left(v_i^2-1\right), \quad i = 1,2.$$

Taking

$$I^*(\theta) = \frac{1}{(1-\theta^2)^2} \le I(\theta) = \frac{1+\theta^2}{(1-\theta^2)^2},$$

it is a routine exercise to check that (2.1) holds. Note that in this case, $\hat{\theta}$ is a function of the normal scores that is different from, but asymptotically equivalent to, the van der Waerden rank correlation coefficient, which is the estimator proposed by Klaassen and Wellner (1997) for this particular model.

Example 2. (the Farlie-Gumbel-Morgenstern system). Copulas in this class are of the form

$$C_\theta(u_1,u_2) = u_1u_2 + \theta u_1u_2(1-u_1)(1-u_2)$$

with $|\theta| < 1$. Setting $v_i = 2u_i - 1$, elementary calculations yield

$$\dot{\ell}_\theta(u_1,u_2) = \frac{v_1v_2}{1+\theta v_1 v_2}, \quad \dot{\ell}_i(u_1,u_2) = \frac{2\theta v_{3-i}}{1+\theta v_1 v_2}$$

and

$$I_{\theta i}(u_i) = -\frac{2}{\theta v_i} - \frac{1}{\theta^2 v_i^2}\log\left(\frac{1-\theta v_i}{1+\theta v_i}\right),$$

whence

$$W_i(u_i) = \sum_{k=1}^{\infty} \frac{\theta^{2k-1}}{2k(2k+1)}\left(v_i^{2k} - \frac{1}{2k+1}\right), \quad i = 1,2.$$

It is then a simple matter to show that for given $\theta \neq 0$, the right-hand side of (2.1) is not proportional to $\dot{\ell}_\theta$.

4. Discussion

Although this remains to be checked in full generality, the nature of condition (2.1) makes it improbable that the omnibus estimator $\hat{\theta}$ is semiparametrically efficient for any of the most common parametric copula models, say within the broad Archimax class recently introduced by Capéraà et al. (2000). Using techniques described in Chapter 3 of Bickel et al. (1993), it would be of interest to determine, for a variety of such models, the extent to which this procedure is inefficient by comparing its limiting variance (already determined by Genest et al. 1995) to the semiparametric information bound.

Alternative semiparametric estimation strategies could also be envisaged. As a step towards the construction of a more efficient procedure, consider the situation in which data $(X_{11}, X_{21}), \ldots, (X_{1n}, X_{2n})$ arise from a bivariate distribution of the form (2.1) with unknown marginals F_1, F_2 and arbitrary, but *known*, absolutely continuous copula C with density c. As before, let $X_{i(j)}$ denote the jth order statistic among the observations of $X_{i1}, \ldots, X_{in}$.

Suppose for a moment that each F_i belonged to a parametric family $(F_{i\alpha_i})$ with corresponding densities $(f_{i\alpha_i})$. Obviously, the α_i's should then be estimated, not from the log-likelihoods of the individual univariate margins

$$L_i(\alpha_i) = \sum_{j=1}^{n} \log f_{i\alpha_i}(X_{ij}),$$

but from the overall log-likelihood

$$L_1(\alpha_1) + L_2(\alpha_2) + \sum_{j=1}^{n} \log c\{F_{1\alpha_1}(X_{1j}), F_{2\alpha_2}(X_{2j})\}.$$

Likewise, nonparametric estimators $\hat{F}_i$ of the F_i's should exploit the additional information provided by the fact that C is known. An alternative to the omnibus procedure leading to $\hat{\theta}$ would then be as follows: (i) derive an initial consistent estimate $\tilde{\theta}$ of the dependence parameter by some simple method, such as the inversion of Kendall's tau; (ii) compute the $\hat{F}_i$'s assuming $C_\theta = C_{\tilde{\theta}}$; (iii) replace F_{in} by $\hat{F}_i$ in the pseudo-log-likelihood (1.3), and maximize it with respect to θ. At present, however, it is not clear how efficient nonparametric estimators $\hat{F}_i$'s might be constructed using knowledge about C. One option might be to observe that among estimators of the form

$$\hat{F}_i(x) = \sum_{j=1}^{n} p_{ij} 1\left(X_{i(j)} \le x\right),$$

with arbitrary positive weights adding up to one, the standard empirical distribution function corresponds to the choice of p_{ij}'s that maximizes the empirical likelihood (Owen 1990)

$$\sum_{j=1}^{n} \log p_{ij} = \sum_{j=1}^{n} \log P\left(X = X_{i(j)}\right),$$

i.e., the likelihood of the data with respect to the counting measure on the order statistics $X_{i(j)}$. When C is known, analogy with the parametric case would then suggest that the p_{ij}'s be chosen to maximize

$$\sum_{j=1}^{n} \log p_{1j} + \sum_{j=1}^{n} \log p_{2j} + \sum_{j=1}^{n} \log c\left(\sum_{k=1}^{R_{1j}} p_{1k}, \sum_{k=1}^{R_{2j}} p_{2k}\right),$$

where R_{ij} denotes the rank of X_{ij} among $X_{i1}, \ldots, X_{in}$. This seems intuitively reasonable, since knowledge of the copula might lead one to assign higher weights to those pairs of observations (or more precisely to the induced relative ranks $(R_{1j}, R_{2j})/n$) that occurred in comparatively more likely regions of the unit square. And indeed, the $\hat{F}_i$'s can be proved to be asymptotically normal and unbiased estimators of the marginal distributions in that case. Although these estimators are possibly more efficient than the F_{in}'s in large samples, extensive Monte Carlo simulations suggest that coupled with an initial estimation of θ by inversion of Kendall's tau, this choice of nonparametric estimator for the marginals does not yield any significant improvement in the estimation of the dependence parameter of the various copula models considered.

Appendix: Proof of the characterization result

From Section 3.4 of Bickel et al. (1993), efficient semiparametric inference must be based on the "efficient score," i.e., the residual of the projection of $\dot{\ell}_\theta$ on a closed linear subspace of functions that are square-integrable and have zero mean with respect to the measure C_θ. In the bivariate case, and under the regularity conditions of Proposition 4 on p. 166 of Bickel et al. (1993), this so-called tangent space comprises all functions $t(u_1, u_2)$ expressible in the form

$$t(u_1, u_2) = \sum_{i=1}^{2} \left\{\gamma_i(u_i) + \dot{\ell}_i(u_1, u_2)\Gamma_i(u_i)\right\} \tag{4.1}$$

in terms of arbitrary functions γ_i with

$$\Gamma_i(u_i) = \int_0^{u_i} \gamma_i(v)dv \quad \text{and} \quad \Gamma_i(1) = 0, \quad i = 1, 2.$$

The same regularity conditions imply that

$$\int_0^1 \dot{\ell}_i(u_1, u_2) c_\theta(u_1, u_2) du_{3-i} = 0,$$

and (through integration by parts)

$$\int_0^1 \gamma_i(u_i) c_\theta(u_1,u_2) du_i = -\int_0^1 \Gamma_i(u_i) \dot{\ell}_i(u_1,u_2) c_\theta(u_1,u_2) du_i.$$

Hence it follows from (4.1) that

$$\gamma_i(u_i) = \int_0^1 t(u_1,u_2) c_\theta(u_1,u_2) du_{3-i} = \mathrm{E}\{t(U_1,U_2) \mid U_i = u_i\}, \tag{4.2}$$

where (U_1,U_2) has distribution C_θ. Thus if $h(u_1,u_2)$ is a zero-mean function that is orthogonal to the tangent space, it must satisfy

$$\mathrm{E}\{h_i(U_1,U_2) \mid U_i = u_i\} = 0, \tag{4.3}$$

where $h_i(u_1,u_2) = \partial h(u_1,u_2)/\partial u_i$, $i = 1,2$, because use of (4.1) and integration by parts shows that $\mathrm{E}\{h(U_1,U_2)t(U_1,U_2)\}$

$$\begin{aligned} &= \sum_{i=1}^{2} \int_0^1 \int_0^1 h(u_1,u_2) \left\{ \frac{\partial}{\partial u_i} \Gamma(u_i) c_\theta(u_1,u_2) \right\} du_i du_{3-i} \\ &= -\sum_{i=1}^{2} \int_0^1 \int_0^1 \Gamma(u_i) c_\theta(u_1,u_2) \left\{ \frac{\partial}{\partial u_i} h(u_1,u_2) \right\} du_i du_{3-i} \\ &= -\sum_{i=1}^{2} \mathrm{E}\left[\Gamma_i(u_i) \mathrm{E}\{h_i(U_1,U_2) \mid U_i = u_i\}\right] = 0, \end{aligned}$$

for all choices of γ_i's.
In particular, the efficient score function must be of the form $h(u_1,u_2) = \dot{\ell}_\theta(u_1,u_2) - t(u_1,u_2)$ for a specific t of the form (4.1) with γ_i's that meet condition (4.3). As shown in Section 4.7 of Bickel et al. (1993), this leads to a pair of coupled Sturm-Liouville differential equations involving the γ's, their primitives and their derivatives, viz.

$$I_{\theta i}(u_i) + \gamma_i'(u_i) - I_{ii}(u_i)\Gamma_i(u_i) + \int_0^1 \dot{\ell}_{12}(u_1,u_2)\Gamma_{3-i}(u_{3-i}) du_{3-i} = 0$$

for $i = 1,2$. These results were used by Klaassen and Wellner (1997) to assess the semiparametric efficiency of the normal scores (or van der Waerden) rank correlation coefficient as an estimator of θ in the normal copula model. The same can be done with the omnibus estimator $\hat{\theta}$, whose asymptotic variance is given in Proposition 2.1 of Genest et al. (1995) as

$$\sigma^2 = \frac{1}{I(\theta)} + \frac{1}{I^2(\theta)} \operatorname{var}\left\{ \sum_{i=1}^{2} W_i(U_i) \right\},$$

where the W_i's are defined as in Section 2 to have zero expectation and derivative $I_{\theta i}$. The score function implicitly used by these authors is

$$h(u_1,u_2) = \frac{1}{\sigma^2 I(\theta)} \left\{ \dot{\ell}_\theta(u_1,u_2) + \sum_{i=1}^{2} W_i(u_i) \right\},$$

which satisfies (4.3) and hence is orthogonal to the tangent space, because the regularity conditions imply that

$$\mathrm{E}\left\{ \dot{\ell}_{\theta i}(U_1,U_2) \mid U_i = u_i \right\} = -I_{\theta i}(u_i), \quad i = 1,2.$$

For $\hat{\theta}$ to be efficient, however,

$$\begin{aligned} t(u_1,u_2) &= \dot{\ell}_\theta(u_1,u_2) - h(u_1,u_2) \\ &= \left\{ 1 - \frac{1}{\sigma^2 I(\theta)} \right\} \dot{\ell}_\theta(u_1,u_2) - \frac{1}{\sigma^2 I(\theta)} \sum_{i=1}^{2} W_i(u_i), \end{aligned}$$

should also be of the form (4.1) with γ_i's given by (4.2), viz.

$$\gamma_i(u_i) = -\frac{1}{\sigma^2 I(\theta)} \left[W_i(u_i) + \mathrm{E}\{ W_{3-i}(U_{3-i}) \mid U_i = u_i \} \right],$$

on account of the fact that $\mathrm{E}\{\dot{\ell}_\theta(U_1,U_2) \mid U_i = u_i\} = 0$ for $i = 1,2$. Accordingly, identity (2.1) must hold with $I^*(\theta) = 1/\sigma^2 \leq I(\theta)$, whence the result.

Acknowledgment

This project was initiated while both authors were at the Université libre de Bruxelles. Partial funding in support of the work was provided by the Natural Sciences and Engineering Research Council of Canada and by the Fonds pour la formation de chercheurs et l'aide à la recherche du Gouvernement du Québec.

References

Bandeen-Roche, K. J. and K.-Y. Liang (1996), Modelling failure-time associations in data with multiple levels of clustering. *Biometrika* **83**, 29–39.

Bickel, P. J., C. A. J. Klaassen, Y. Ritov and J. A. Wellner (1993), *Efficient and Adaptive Estimation for Semiparametric Models*. Baltimore, MD: The Johns Hopkins University Press.

Capéraà, P., A.-L. Fougères and C. Genest (2000), Bivariate distributions with given extreme value attractor. *J. Multivariate Anal*. **72**, 30–49.

Clayton, D. G. and J. Cuzick (1985), Multivariate generalizations of the proportional hazards model (with discussion). *J. R. Statist. Soc*. A **148**, 82–117.

Genest, C. (1987), Frank's family of bivariate distributions. *Biometrika* **74**, 549–555.

Genest, C., K. Ghoudi and L.-P. Rivest (1995), A semiparametric estimation procedure of dependence parameters in multivariate families of distributions. *Biometrika* **82**, 543–552.

Hougaard, P. (1989), Fitting a multivariate failure time distribution. *IEEE Trans. Reliab.* **38**, 444–448.

Joe, H. (1997), *Multivariate Models and Dependence Concepts*. London: Chapman and Hall.

Klaassen, C. A. J. and J. A. Wellner (1997), Efficient estimation in the bivariate normal copula model: normal margins are least favourable. *Bernoulli* **3**, 55–77.

Klugman, S. A. and R. Parsa (1999), Fitting bivariate loss distributions with copulas. *Ins. Math. Econ.* **24**, 139–148.

Mikusiński, P., H. Sherwood and M. D. Taylor (1992), Shuffles of min. *Stoch.* **13**, 61–74.

Oakes, D. (1982), A model for association in bivariate survival data. *J. R. Statist. Soc*. B **44**, 414–422.

Oakes, D. (1986), Semiparametric inference in a model for association in bivariate survival data. *Biometrika* **73**, 353–361.

Oakes, D. (1989), Bivariate survival models induced by frailties. *J. Amer. Statist. Assoc.* **84**, 487–493.

Oakes, D. (1994), Multivariate survival distributions. *J. Nonparametr. Statist.* **3**, 343–354.

Owen, A. B. (1990), Empirical likelihood ratio confidence regions. *Ann. Statist.* **18**, 90–120.

Shih, J. H. and T. A. Louis (1995), Inferences on the association parameter in copula models for bivariate survival data. *Biometrics* **51**, 1384–1399.

Zheng, M. and J. P. Klein (1995), Estimates of marginal survival for dependent competing risks based on an assumed copula. *Biometrika* **82**, 127–138.