

THE ROBUSTNESS OF MAINTENANCE OPTIMIZATION TO MODELLING ASSUMPTIONS

Cornel Bunea

Delft University of Technology, Delft, e-mail: c.bunea@its.tudelft.nl

Tim Bedford

Strathclyde Business School, Glasgow, e-mail: Tim@mansci.strath.ac.uk

ABSTRACT

Maintenance optimization requires the use of many modelling assumptions. We focus on issues related to data interpretation, and in particular to competing risk. Competing risk models are used to interpret data and estimate some of the inputs needed for maintenance optimization.

1 INTRODUCTION

Independent competing risks models have been studied for some time. Here we discuss key issues of sensitivity to competing risk model assumptions, in particular the choice of copula to “remove” the effect of censoring from censored data. By observing independent samples of competing risks we can estimate the sub-survival functions. Assuming independence of competing risks we can identify the underlying marginal distributions. The assumption of independence can not always be taken when failures are censored by preventive maintenance. The assumption of independence would imply that maintenance engineers take no account of the state of the component when taking the decision to preventively maintain. It is more reasonable to make a dependence assumption between censoring processes. We assume that the dependence between risks is given by a copula. Under this assumption, Zheng and Klein [10] showed that the observable data determines a unique model. In this paper we will model the joint distribution in question by three different types of copula, we will apply the results on the problem of age replacement policy and we will show that the optimal replacement time is not dependent of the choice of copula. In Section 1 we present some basics about competing risks problems and a brief introduction to maintenance policies. In Section 2 the results of Zheng and Klein concerning the copula solution of the competing risks models are presented. In next sections we present the results of Bunea and Bedford [4] and we introduce four families of copula and some of their properties. In the end of the study a numerical example is given and a short discussion will succeed.

2 COMPETING RISK AND MAINTENANCE OPTIMIZATION

2.1 Competing Risk

In the competing risk approach we model the data as a renewal process, that is as a sequence of i.i.d. variables Z_1, Z_2, \dots . Each observable Z is the minimum of two variables. The lifetime of the component is X : this is the lifetime that the component would reach if it were not preventively maintained. The preventive maintenance (PM) time of the component is Y : this is the time at which the component would be preventively maintained if it didn't fail first. Clearly,

$$Z \equiv \min(X, Y).$$

(In fact, usually X will be the minimum of several variables giving the time to failure by a particular failure mode: we shall just consider the case of one failure mode.) The observable data will allow us to estimate the *sub-survival functions*,

$$S_X^*(t) \equiv Pr\{X > t, X < Y\} \quad \text{and} \quad S_Y^*(t) \equiv Pr\{Y > t, Y < X\},$$

but not the true survivor functions of X and Y . Hence we are not able to estimate the underlying failure distribution for X without making additional, non-testable, model assumptions. A characterization of those distributions for X that are possible for given sub-survival functions was made in [3].

By specifying a copula for the underlying joint distribution of X and Y one *can* identify the marginals (and the full joint distribution) [10]. However the choice of such a copula is difficult to make: Bedford [2] suggests doing this by specifying the Spearman's rank correlation between X and Y , and then using the copula with minimum information with respect to the independent copula (that is, the "most independent" copula with the given Spearman rank correlation).

2.2 Maintenance Optimization

To keep things simple we just consider the *age replacement policies*. Recall that an age replacement policy is one for which replacement occurs at failure or at age θ , whichever occurs first. Unless otherwise specified, θ is taken to be a constant.

In the finite time span replacement model we will try to minimize expected cost $C(\theta)$ experienced during $[0, \theta]$, where cost may be computed in money units, time, or some appropriate combination. For a infinite time span, an appropriate objective function is expected cost per unit of time, expressed as

$$\gamma(\theta) \equiv \lim_{\theta \rightarrow \infty} \frac{C(\theta)}{\theta}.$$

Letting $N_1(\theta)$ denote the number of failures during $[0, \theta]$ and $N_2(\theta)$ denote the number of planned preventive maintenances during $[0, \theta]$, we may express the expected cost during $[0, \theta]$ as

$$C(\theta) \equiv c_1 E\{N_1(\theta)\} + c_2 E\{N_2(\theta)\},$$

where c_1 is the cost of critical failure and c_2 is the cost for planned replacement. We only consider non-random age replacement in seeking the policy minimizing the specific cost $\gamma(\theta)$ for an infinite time span.

Starting from the definition of the specific cost

$$\gamma(\theta) \equiv \lim_{\theta \rightarrow \infty} \left[c_1 \frac{E\{N_1(\theta)\}}{\theta} + c_2 \frac{E\{N_2(\theta)\}}{\theta} \right]$$

[1] showed that

$$\gamma(\theta) \equiv \frac{c_1 F(\theta) + c_2 S(\theta)}{\int_0^\theta S(t) dt}.$$

Then $\gamma(0) = \infty$ and $\gamma(\infty) = c_1 / \int_0^\theta S(t) dt$. Differentiating γ to find the optimum, $\frac{d\gamma(\theta)}{d\theta} = 0$, we obtain the equation

$$r(\theta) \int_0^\theta S(t) dt - F(\theta) = \frac{c_2}{c_1 - c_2}.$$

When $F_X(x)$ has an increasing failure rate, the optimal replacement time θ_0 is the unique solution of the above equation. For a r.v. with constant failure rate or decreasing failure rate the specific cost has not an optimum ($sign(\frac{d\gamma(\theta)}{d\theta})$ is constant), thus this type of maintenance policy is not appropriate for a such r.v.

When we have as primary parameter Kendall's tau and the information over τ is given by a distribution function $F_\tau(\tau)$ with density $f_\tau(\tau)$, the specific cost is dependent on τ and θ :

$$\gamma(\tau, \theta) \equiv \frac{c_1 F(\tau, \theta) + c_2 S(\tau, \theta)}{\int_0^\theta S(\tau, t) dt}.$$

So the long term specific cost given θ is

$$\gamma(\theta) = \int_0^\theta C(\tau, \theta) f_\tau(\tau) d\tau$$

and the optimal replacement time θ_0 is obtaining minimizing $\gamma(\theta)$.

If as a primary parameter we take Spearman's rho, analogous formulas as in Kendall's tau case are obtained.

3 ZHENG AND KLEIN COPULA SOLUTION TO COMPETING RISKS MODELS

As defined by Schweizer and Wolff [9] the copula of two random variables X and Y is the distribution C on the unit square $[0, 1]^2$ of the pair $(F_X(X), F_Y(Y))$ (recall that for a continuous random variable X with pdf F_X , the random variable $F_X(x)$ is always uniformly distributed on $[0, 1]$). The functional form of $C : [0, 1]^2 \rightarrow \mathbb{R}$ is

$$C(u, v) \equiv H(F_X^{-1}(u), F_Y^{-1}(v)),$$

where H is the joint distribution function of (X, Y) and F_X^{-1} and F_Y^{-1} are the right-continuous inverses of F_X and F_Y . Under independence of X and Y the copula is $C(u, v) = uv \equiv \Pi$, and any copula must fall between $M(u, v) \equiv \min(u, v)$ and $W(u, v) \equiv \max(u + v - 1, 0)$, the copulas of the upper and lower Fréchet bounds [8]. Under the assumption of

independence of X and Y , the marginal distribution functions of X and Y are uniquely determined by the sub-survival functions of X and Y . Zheng and Klein [10] showed the more general result that, if the copula of (X, Y) is known, then the marginal distributions functions of X and Y are uniquely determined by the competing risk data. This result is captured in the following theorem:

Theorem 1 *Suppose the marginal distribution functions of (X, Y) are continuous and strictly increasing in $(0, \infty)$. Suppose the copula C is known and the corresponding probability measure for any open set of the unit square is positive. Then F_X and F_Y , the marginal distribution functions of X and Y , are uniquely determined by the subdistribution functions.*

We show briefly why the marginals are identifiable in the case that densities and subdensities exists. By definition we have that the subdistribution function of X is

$$F_X^*(t) \equiv P\{X \leq t, X < Y\}.$$

After a straightforward calculation we get:

$$\begin{aligned} F_X^*(t) &= \int_0^t \int_x^\infty h(x, y) dx dy = \int_0^t H_X(x, \infty) - H_X(x, x) dx = \\ &= F_X(t) - \int_0^t H_X(x, x) dx = F_X(t) - \int_0^t C_u(F_X(x), F_Y(x)) f_X(x) dx, \end{aligned}$$

where $h(x, y)$ is the joint density function of X and Y and $H_X(x, \infty)$ respectively $H_X(x, x)$ denote the first order partial derivative $\frac{\partial}{\partial x} H(x, y)$ calculated in (x, ∞) respectively in (x, x) . We obtain an analogous formula for F_Y^* . From this formula it follows that the marginal distributions functions F_X and F_Y are solutions of the following system of ordinary differential equations:

$$\begin{cases} \{1 - C_u(F_X(t), F_Y(t))\} F_X'(t) = F_X^{*'}(t) \\ \{1 - C_v(F_X(t), F_Y(t))\} F_Y'(t) = F_Y^{*'}(t) \end{cases}$$

with initial conditions $F_X(0) = F_Y(0) = 0$, where $C_u(F_X(t), F_Y(t))$ and $C_v(F_X(t), F_Y(t))$ denote the first order partial derivatives $\frac{\partial}{\partial u} C(u, v)$ and $\frac{\partial}{\partial v} C(u, v)$ calculated in $(F_X(t), F_Y(t))$.

4 THREE FAMILIES OF COPULAE

Work of Zheng and Klein [10] suggests that the important factor for an estimate of the marginal survival function is a reasonable guess at the strength of the association between competing risks and not the functional form of the copula. For this reason we chose [4] Archimedean copula with which it is easy to work from the mathematical point of view and we calculated for one family of Archimedean copula the average specific cost with optimal replacement time. Now we introduce other two classes of copula and another family of Archimedean copula.

Minimally informative copula

This copula was originally introduced to knowledge dependence in uncertainty analysis [5]. Minimally informative copula has minimal information (taken with respect to the uniform distribution) with the given correlation among all candidate distributions. The relative information function of a continuous distribution function with density $f_X(x)$ with respect to the continuous distribution function with density $f_Y(y)$ is:

$$I(f_X|f_Y) = \int \log\left(\frac{f_X(t)}{f_Y(t)}\right)f_X(t)dt.$$

Then the information function of minimally informative copula taken with respect to the uniform distribution is:

$$I(c|u) = \int \int_{I^2} c(x, y) \log\{c(x, y)\}dxdy,$$

where c is the copula density. Meeuwissen and Bedford [7] showed how the minimally informative copula with given correlation coefficient could be determined. This copula has density of the form:

$$c(x, y) = k(x, \theta)k(y, \theta) \exp^{\theta(x-0.5)(y-0.5)},$$

where $\theta = \theta(\rho)$ is a certain monotone increasing function of the correlation coefficient ρ , and the function $k(\cdot, \theta)$ is determined as the solution to an integral equation:

$$k(x, \theta) = \left[\int_0^1 k(y, \theta) \exp^{\theta(x-0.5)(y-0.5)} dy \right]^{-1}.$$

Bedford proposed in 1997 a discrete approximation to the density of minimally informative copula. This method is based on a DAD algorithm.

Diagonal band distribution

Definition 1 For positive measures of association the density of the diagonal band distribution, $b_\alpha(x, y)$ has a mass in a band around the diagonal $y = x$. The bandwidth β equals $1 - \alpha$; i.e.

$$P\{X - \beta < Y < X + \beta\} = 1.$$

The density $b_\alpha(x, y)$ takes values 0, $1/2\beta$ and $1/\beta$ on the five regions it divides the unit square into,

$$b_\alpha(x, y) = \frac{1/2}{1 - \alpha} (1_{\{\alpha-1 \leq x-y \leq 1-\alpha\}} + 1_{\{1-x-y \geq \alpha\}} + 1_{\{1-x-y \leq -\alpha\}}),$$

with $0 \leq \alpha \leq 1$ and $0 \leq x, y \leq 1$.

Zheng and Klein assumed continuous density to get identifiability, so just using diagonal band distribution is not possible. Hence, we will consider mixtures of diagonal band distribution with relative information with respect to the uniform density given a correlation.

Definition 2 Let $0 \leq p \leq 1$. A probability distribution $M(\alpha)$, $M : [-1, 1] \rightarrow [0, 1]$ is called a mixing function if its probability density function consists of an absolutely continuous part $m(\alpha) \geq 0$ with $\int_{-1}^1 m(\alpha) d\alpha = p$ and a discrete part with atoms $p_j > 0$ at α_j , $-1 \leq \alpha_{j-1} \leq \alpha_j \leq 1$, $\cup_j \{\alpha_j\} = \mathcal{A}$ and $\sum_j p_j = 1 - p$. \mathcal{A} may be the empty set.

Definition 3 Let $M(\alpha)$ be a mixing function. A mixture $f_M(x, y)$ of diagonal band densities $b_\alpha(x, y)$ is defined as

$$f_M(x, y) = \int_{-1}^1 b_\alpha(x, y) dM(\alpha).$$

Meeuwissen [6] proved that mixtures f_M with correlation ρ have less relative information with respect to the uniform distribution $I(f_M|u)$ than the diagonal band density with the same correlation. This is another reason to consider mixture of diagonal band distributions to model an incompletely specified joint probability distribution. Using the results of Meeuwissen we can approximate the mixing densities that give mixture with minimal relative information with respect to the uniform density very well with a mixture of the uniform density and a beta density. Thus we can determine directly that mixing measure in the class of beta densities that gives a mixture of diagonal band densities with minimal relative information with respect to the uniform density given correlation.

Archimedean copula

We considered in the previous work only the Gumbel family of copulae:

$$C_\alpha(u, v) = \exp(-[(\ln \frac{1}{u})^\alpha + (\ln \frac{1}{v})^\alpha]^\frac{1}{\alpha})$$

$$\varphi_\alpha(t) = (-\ln t)^\alpha \quad \text{with } \alpha \in [1, \infty) \quad \text{and } C_1 = \Pi, C_\infty = M.$$

For this study we will take into account another family of Archimedean copula.

$$C_\alpha(u, v) = \exp(1 - [(1 - \ln u)^\alpha + (1 - \ln v)^\alpha - 1]^\frac{1}{\alpha})$$

$$\varphi_\alpha(t) = (1 - \ln t)^\alpha - 1 \quad \text{with } \alpha \in [0, \infty) \quad \text{and } C_1 = \Pi, C_\infty = M.$$

Using Theorem 2 [4], we obtain the following relation between Kendall's tau and the parameter α :

$$\tau(\alpha) = \exp(2) * \text{gamma}(2 - \tau, 2) * 2^\tau + \tau - 3) / \tau.$$

5 MAINTENANCE OPTIMIZATION FOR DIFFERENT FAMILIES OF COPULAE

To study the robustness with respect to choice of copula, we will consider four sets of numerical experiments, one set for each type of copula presented above.

Since the costs of critical failure can be much higher than those of planned maintenance (because of other consequences to the system beyond the need simply to replace the

failed unit), we assume that c_1 is much larger than c_2 . Since actual plant data shows a considerable number of preventive maintenance actions we assume that $p = Pr(X < Y)$ is small. Specifically we take $c_1/c_2 = 10$, $p = 0.3$. Let:

$$S_X^*(t) = 0.3t^{1.5} \exp(-0.4t^{2.5})$$

and

$$S_Y^*(t) = 0.7t^{1.5} \exp(-3.2t^{2.5}),$$

be the sub-survival functions of X and Y . In the previous study we considered Kendall's tau as the measure of association for the pair (X, Y) . Now we will consider as a primary parameter Spearman's rho for the minimally informative copula and mixture of diagonal band, and Kendall's tau for the others copulae. To obtain a distribution for Kendall's tau we ask an expert to give quantiles for the probability q defined in the previous section. If the expert gives 5% and 95% quantiles then we can fit a beta distribution. Specifically if $Pr\{q \leq 0.7\} = 0.05$ and $Pr\{q \leq 0.95\} = 0.95$, then the 5% and 95% quantiles for τ are $Pr\{\tau \leq 0.4\} = 0.05$ and $P\{\tau \leq 0.9\} = 0.95$. Taking the beta distribution as appropriate for τ , we easily obtain the parameters of this distribution a and b given the above quantiles by Newton's method as: $a = 5.6705$ and $b = 2.7322$. For Spearman's rho we will consider the uniform distribution on $[0, 1]$. The average specific cost with optimal replacement time for different copula are shown in Figure 6.

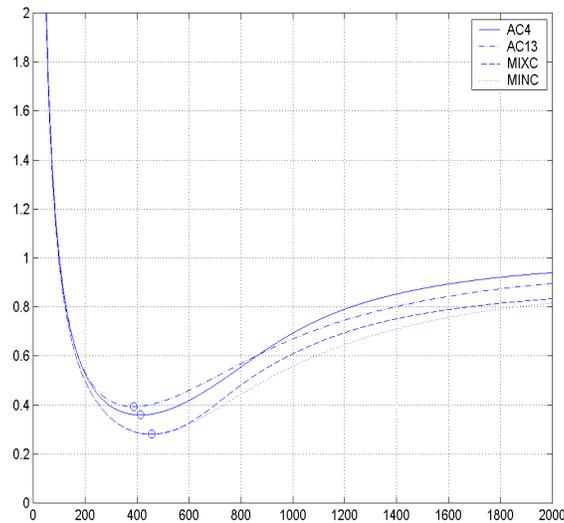


Figure 1. The average specific cost with optimal replacement time. AC4-the result for Gumbel copula; AC13 - the result for the other family of Archimedean copula; MIXC - the result for the mixture of diagonal band distribution; MINC - the result for minimally informative copula

6 CONCLUSIONS

The results show little difference among the four different families of copula, regarding the optimal replacement time. Figure 6 suggests that the important factor for determine the optimal replacement time is the measure of association between risks X and Y and not the functional form of the copula. This results confirm Zheng and Klein study over the robustness of copula-graphic estimator.

The results also show that the optimal replacement times for all families of copula studied can be found in a small interval. This is most important since it suggests that we can robustly determine the optimal replacement time even without being certain about the optimal costs.

References

- [1] R.E. Barlow, F. Proschan, *Mathematical theory of reliability*, 1965, Wiley, New York.
- [2] T. Bedford , “On the use of minimally informative copulae in competing risk problems”, *Statistical Probabilistic Models in Reliability*, C. Ionescu and N. Limnios (eds), 1998, Birkhauser, Boston.
- [3] T. Bedford, I. Meilijson, “A characterization of marginal distributions of (possibly dependent) lifetime variables which right censor each other”, *Annals of Statistics*, **25**, 1997, pp 1622-1645.
- [4] C. Bunea, T. Bedford, “The effect of model uncertainty on maintenance optimization”, due to appear in *IEEE*
- [5] R.M. Cooke, “Unicorn: Methods and node for uncertainty analysis” *AEA Technology (for the European Safety and Reliability Association)*, 1995
- [6] A.M.H. Meeuwissen, “Diagonal band probability distributions”, *PhD Thesis*,
- [7] A.M.H. Meeuwissen, T. Bedford, “Minimally informative distribution with given rank correlation for use in uncertainty analysis”, *Journal of Statistical Computation and Simulation*
- [8] R.B. Nelsen, *An Introduction to copulas*, 1995, Springer
- [9] B. Schweizer, and E.F. Wolff, “On nonparametric measures of dependence for random variables” *Annals of Statistics*, **9**, 1981, pp 879-885.
- [10] M. Zheng, J.P. Klein, “Estimates of marginal survival for dependent competing risks based on an assumed copula”, *Biometrika*, vol **82** , 1995, pp 127-138.