Lecture Notes in Management Science (2013) Vol. **5**: 33–38
5th International Conference on Applied Operational Research, Proceedings

ISSN 2008-0050 (Print), ISSN 1927-0097 (Online)

On the use of expert opinion to characterise the joint behaviour of competing risks in industrial accelerated life testing

Herbert Hove

School of Statistics and Actuarial Science,
University of the Witwatersrand, Johannesburg, South Africa
herbert.hove@wits.ac.za

Abstract. Distribution identifiability issues arise quite naturally through competing risks in reliability. In particular, modeling and analysis of recurrent events which include the impact maintenance has on the lifetime distribution of repairable systems is an interesting practical application. This paper discusses the problem of modeling the joint behaviour of condition-based preventive maintenance (PM) and corrective maintenance (CM) in the competing risk context using copulas. Specifically, how expert judgement and accelerated life testing data can be used to estimate the copula dependence parameter and quantify its uncertainty through Monte-Carlo simulations is discussed.

Keywords: competing risk; identifiability; copula; expert opinion; simulations

Introduction

Competing risks in industrial accelerated life testing

Systems generally fail from different causes and these are often present in industrial accelerated life testing (ALT). When cause and lifetime information is available, competing risk theory provides the appropriate model for analyzing failure data. The presence of competing failure modes in industrial ALT has been studied for some time, see e.g McCool (1978), Klein and Basu (1981, 1982a, 1982b). Nelson (1990) contains an entire chapter on competing failure modes in ALT.

The interpretation and analysis of data concerning competing failure causes in engineering, biostatistical and related application fields remains an open problem. Denote by J the cause of system failure for $J = 1,...,d$. In a typical competing risk situation, these d failure causes are competing to be responsible for terminating the system's service at each stress level. What is observed is the time to first failure $T \geq 0$ (possibly censored) and the failure cause J. If T is censored, J will remain unknown. Hence only the pair (T, J) is in general observable, called the identified minimum (Heckman and Honore, 1989; Crowder, 1994).

Depending on the application a vector of covariates $z = (z_1,..., z_p)$ recording study units characteristics, stress levels etc. may also be observed.

Theoretical approaches to the competing risk problem introduce hypothetical or latent failure time random variables $Y_1,...,Y_d$, $0 \leq Y_j < \infty$ representing system failure times corresponding to the d failure causes $J = 1,...,d$. In actual study conditions when all risks are acting, only the earliest of these theoretical failure times $T = \min(Y_1,...,Y_d)$ is observed along with J, the cause of failure. Hence T is formally the lifetime of a series or weakest link system having d components where system failures are a result of a unique failing component. As used here, series system refers to how system failure depends on component failure, not physical connections of components.

Under the latent failure time approach, competing risk problems concerning interrelations among failure causes and the effects of cause removal are formulated in terms of the Y_j's and the joint survival (multiple decrement) function

$$S(t_1,...,t_d) = \left[P \bigcap_{j=1}^{d} \left(Y_j > t_j \right) \right]$$

which is considered absolutely continuous and satisfying $S(0,...,0) = 1$ and $S(\infty,...,\infty) = 0$ where each $t_j \geq 0$. This model is necessary whenever study questions are posed in terms of the (marginal) survival distribution of failure time from a specific cause when others are removed. This is particularly so in reliability applications where it is often interesting to ask the question: "*What would be the effect of eliminating a failure cause on the reliability specification of the system*?" Cause removal can in practice be achieved through system redesign for example.

Though a formal mathematical model, the latent failure time model in general presents problems of interpretation and data analysis. It is however of value in situations where random censorship is modelled as a competing risk arising from a mechanism external to the failure process under consideration (Prentice et. al., 1978). A typical example in reliability (see eg. Cooke, 1996) is the effect maintenance has on the basic failure time distribution of a repairable system as follows. Denote by X_2 the random time to system failure if no preventive maintenance (PM) were performed. System failure can be avoided by a potential condition-based PM action that can occur at random time X_1. Thus at each stress level, the maintained system can be taken out of service by corrective maintenance (CM) or PM which are the two competing risk variables. The corresponding latent failure time random variables X_1 and X_2 represent times to a maintenance action.

This probability structure allows condition-based PM (random censoring) to be one of the causes responsible for taking the system out of service. It is a unique competing risk formulation in that it guarantees an underlying system failure mechanism that is not influenced by the presents of random censoring. Besides attaching physical meaning to a latent failure time to a CM action, it also guarantees that the marginal distribution arising from the removal of censoring is the basic failure time distribution of the system. Undoubtedly, the system failure time distribution is the relevant target of estimation from which reliability metrics are derived. Hence, the aim of this paper is to address the problem of estimating the basic failure time distribution for a stochastically and continuously deteriorating repairable system with the maintenance effect (a competing risk) removed.

Modeling the CM - PM process in a competing risk framework

A well-known difficulty with the latent failure time approach (Tsiatis, 1975) is that neither the joint nor the marginal survival functions of the competing risk variables are in general identifiable unless if the risks are independent. Since CM and PM are linked through the degradation process of the system, they cannot be taken to be independent. Rather, they are assumed to have a basic dependence structure at each stress level and hence special and more complicated statistical methods are required.

An obvious approach that allows for dependent competing risks is to place parametric restrictions on the joint survival function $S(t_1, t_2)$ in order to study interrelation more generally. Within such parametric models, parameters that describe possible dependencies between latent failure times X_1 and X_2 may be estimated. Crucially however, there must be external evidence to justify the assumed parametric model since dependence arises from a model assumption that cannot be tested by the competing risk data alone. Otherwise a non-zero value of the estimated dependence parameter within such parametric models is not necessarily an indication of dependence between the competing risk variables.

The problem with placing parametric restrictions on $S(t_1, t_2)$ is that the observable competing risks data do not allow one to distinguish between the assumed model and one with independent risks. Hence in addition to the uncertainty due to sampling error, there is also the extra problem of model uncertainty. Other approaches (see e.g. David & Moeschberger, 1978; Meeker, Escobar & Hong, 2009) collapse several related failure modes into fewer groups such that the resulting failure modes are approximately independent and hence identifiable. Depending on the application, interest may be in estimating the subsurvival functions only (Crowder, 2001) which are identifiable. In this study however, identifiability issues remain since the problem is to infer the marginal survival functions when dependent competing risks are acting.

The Copula-Dependent Competing Risks Model in ALT

When in service, degrading systems tend to deliver some kind of signal which is a warning that a failure is approaching. If the signal is detected, a PM is performed to avoid or postpone system failure which is generally perceived to be more costly than degraded or non-critical failure. To be efficient however, PM has to be performed just before a failure occurs. Hence the dependence structure between CM and PM is the degree to which the occurrence of high (low) values for the one variable impacts on the probability of occurrence of values of the other variable. This notion of dependence is a matter of relative ranks and is therefore completely based on copulas.

Estimation of the copula dependence parameter(s) from competing risk data is in general a difficult problem since a set of n pair-wise observations is required. Rather, what is observed is the random variable $Z = \min(X_1, X_2)$ together with the identity of the failure mode that achieved the minimum. Hence, the analysis inevitably has to rely on expert judgments. The use of expert judgement within reliability and science where it is increasing recognised as just another type of scientific data is not new (see e.g. van Noortwijk *et al* 1992; Kurowicka and Cooke 2006) and the numerous references therein. The drawback of this approach is that eliciting dependence from experts in situations where no 'hard data' are available on the variables of interest is a difficult judgmental task.

A unique copula associated with a pair (X_1, X_2) is invariant under strictly increasing transformations of the marginals. Since the dependence between X_1 and X_2 is characterised by this copula, a faithful measure of dependence also needs to exhibit the same invariance property. One such measure is the rank correlation which measures the degree of monotone relationships between variables. The best known rank based distribution-free measures of association are Spearman's rho and Kendall's τ. In terms of the copula function (Carriere, 1994; Nelsen, 1995), they are given by

$$\rho_{X_1X_2} = 12\iint_{I^2} C(u_1,u_2)du_1du_2 - 3 \text{ and}$$

$$\tau_{X_1X_2} = 4\iint_{I^2} C(u_1,u_2)dC(u_1,u_2) - 1 \text{ respectively.}$$

Parameterization of families of copulae by the rank correlation implies that the rank correlation can be taken as the primary parameter. Admittedly, the rank correlation is a numerical quantity with an infinite number of possible values. Thus, judgements about the likelihood and uncertainty of the rank correlation value can be properly expressed through an elicited probability distribution. But the expert can only make specific judgements about certain summaries of the distribution, usually the mean or a number of percentiles. Constructing a fully specified continuous probability distribution from a finite number of these specific judgements remains an ill-posed problem since many other possible distributions would have fitted the elicited judgements equally well. Even if multiple experts were available, they will most likely have different degrees of belief. Hence though independent, the elicited summaries may not be identically distributed.

Contribution of the paper

This paper contributes to the literature by showing through a Monte Carlo simulation study how expert opinion on observable quantities and ALT data can be used to characterise the joint behaviour of competing risks at design stress. The practical reason for eliciting on observable quantities only is that experts are more comfortable with answering questions on observables. The rank correlation, which is the target variable in this study is clearly not an observable quantity. Hence, elicitation variables are required to indirectly infer the appropriate correlation number.

Spearman's rho is a widely used measure of rank correlation largely because its computation is very simple i.e. it is Pearson's product moment correlation computed on the ranks assigned to the observations and average ranks. The major disadvantage of Spearman's rho is that it has no simple direct interpretation in terms of probabilities and is thus difficult to quantify. Though usually considered more difficult to compute than Spearman's rho, Kendall's τ , an alternative rank correlation does have a simple and direct interpretation in terms of probabilities of observing concordance and discordance pairs (Conover, 1999). While uncertainty manifests itself in various ways (Booker & Ross, 2011), probability theory is widely accepted as the leading theory when measuring uncertain quantities or quantifying the uncertainty. Because of its interpretation in terms of observables, Kendall's τ (rank correlation henceforth) is preferred to Spearman's rho in the simulation study.

Assuming a large number of test units from the same population at each stress level, then Z_i, $i = 1, 2, ...$ are independent copies of Z . As is often the case in practice, this paper considers the case where only a few prototype systems are available for testing. Sufficient failure data in ALT are therefore obtained by renewing failed systems and testing continuously. Hence the process $\{Z_1, Z_2, ...\}$ defines a renewal process associated with Z . To yield the right data structure, latent failure time random variables X_1 and X_2 for a repairable system at each stress level are simulated from a model tailored for cases where PM censor critical failure. Using expert judgement, the rank correlation value (and hence the copula dependence parameter) is obtained from simulated data. In the case of the Clayton copula for example, the copula dependence parameter θ is given by $\theta = \frac{2\tau}{1-\tau}$.

The simulations are repeated a number of times yielding several rank correlation values from which few summaries of the expert's distribution are inferred. The uncertainty in Kendall's τ is modeled by fitting a subjective distribution to the elicited distribution summaries from the expert. The rank correlation (and hence the copula parameter) is estimated by a specified percentile of the fitted distribution. Given the fitted copula model and a competing risk sample at each stress level, numerical methods are used to infer the marginal distribution of X_2 at design stress which is of interest at design stress.

References

McCool JI (1978). Competing risk and multiple comparison analysis for bearing fatigue tests. ASLE Trans. 21: 271-284

Klein JP and Basu AP (1981). Weibull accelerated life tests when there are competing causes of failure. Comm. Statist. Theory Methods A10: 2073-2100

Klein JP and Basu AP (1982a). Accelerated life testing under competing exponential failure distribution. IAPQR Trans. 7: 1-20

Klein JP and Basu AP (1982b). Accelerated life tests under competing Weibull causes of failure. Comm. Statisti. Theory Methods A11: 2271-2286

Nelson W (1990). Accelerated testing: Statistical models, Test Plans, and Data Analysis. Wiley Series in Probability and Mathematical Statistics. Wiley, New York

Heckman JJ and Honore BE (1989). The identifiability of the competing risks model. Biometrika 76(2): 325–330

Crowder M (1994). Identifiability crisis in competing risks. International Statistical Review 62(3): 379–391

Prentice R, Kalbfleisch J, Peterson A, Flournoy N, Farewell V and Breslow N (1978). The analysis of failure times in the presence of competing risks. Biometrics 34: 541–554

Cooke RM (1994). The design of reliability data bases, Part I and II. Reliability Engineering and System Safety 51: 137–146, 209–223

Tsiatis A (1975). A nonidentifiability aspect of the problem of competing risks. Proceedings of National Academy of Sciences USA 72(1): 20–22

David HA and Moeschberger ML (1978). The theory of competing risks, High Wycombe: Griffin

Meeker WQ, Escober LA and Hong Y (2009). Using accelerated test results to predict product field reliability. International Statistical Review 61(1): 147–168

Crowder M (2001). Classical Competing Risks. Chapman & Hall/CRC

van Noortwijk JM, Dekker R, Cooke RM and Mazzuchi TA (1992). Expert Judgment in Maintenance Optimization. IEEE TRANSACTIONS ON RELIABILITY 41(3): 427–432

Kurowicka D and Cooke R (2006). UNCERTAINTY ANALYSIS with High Dimensional Dependence Modeling. Wiley Series in Probability and Statistics. John Wiley & Sons

Booker JM and Ross TJ (2011). An evolution of uncertainty assessment and quantification. Scientia Iranica D 18(3): 669–676

Carriere JF (1994). Dependent Decrement Theory. Transactions of Society of Actuaries 46: 45–65

Nelsen RB (1995). An Introduction to Copulas, Springer

Conover WJ (1999). Practical nonparametric statistic. John Wiley and Sons, New York.