isid/ms/2007/06
April 26, 2007
http://www.isid.ac.in/~statmath/eprints

Two component load sharing systems with applications to Biology

J. V. Deshpande
Isha Dewan
U. V. Naik-Nimbalkar

Indian Statistical Institute, Delhi Centre
7, SJSS Marg, New Delhi–110 016, India

Two component load sharing systems with applications to Biology

J. V. Deshpande[1,3], Isha Dewan[2] and U. V. Naik-Nimbalkar[1]
[1]Department of Statistics, University of Pune, Pune
[2] Indian Statistical Institute, New Delhi
[3] Norwegian University of Science and Technology, Trondheim

Abstract

Life times of load sharing parallel systems have been considered in the statistical literature at least since Daniels (1945). The main characteristic of such a two component system is that after the failure of one component the surviving component has to shoulder extra load and hence is prone to failure at an earlier time than what is expected under the original model. In other situations, the failure of one component may release extra resources to the other, thus delaying the system failure. Gross et al (1971) observed that similar considerations affect the functioning of a two organ system. In this paper we first consider several observations schemes and identifiability issues under them. Then we construct a general semiparametric bivariate family of distributions which explicitly models this phenomenon through proportional conditional hazards. McCool (2006) has suggested a test for the hypothesis that the failures take place according to the original model against the alternative hypothesis that the second failure takes place earlier than warranted within the Weibull model. We propose nonparametric tests for the same problem which may be used for any continuous distribution for the component life times. We obtain estimates of the power of the test and observe that it is quite high even for moderately distant alternatives. The tests are applied to several real data sets to illustrate their use.

Key words: Bivariate distributions, censoring, conditional distribution, early failures, order statistics, proportional hazards, semiparametric fam-

ily, sharing resources.

1 Introduction, literature survey and summary

A two component parallel system operates as long as at least one of the components is functioning. Let us denote the life of the components 1 and 2 by random variables U_1 and U_2, respectively. Then the life of the system is given by the random variable $Y = max(U_1, U_2)$. Because of the nature of the system, it continues to function even after the failure of one of the components. However, failure of one component can possibly put additional load on the surviving component and hence affect its functioning and hence the functioning of the system. This may result in stochastic reduction of the residual life time.

Gross et al (1971) observed that two organ subsystems in human body typically show this pattern. If one organ fails, the surviving organ is usually subject to higher failure rate. If a patient gets his kidney removed due to to some illness, then the second kidney shows a higher failure rate. However, if a kidney is removed because of an accident, then the second kidney may not exhibit an increased failure rate. The authors develop a survival distribution for such two organ systems . They assume that if one organ fails, then the other organ has a higher failure rate. However, both failure rates are assumed to be constant in time. The parameters of the proposed distribution are estimated iteratively.

The earliest work on load sharing models is due to Daniels (1945)and Rosen (1964). They observed that yarns and cables in a bundle fail only when the last fibre (or wire) in the bundle breaks. A bundle of fibres can be considered as a parallel system subject to a constant tensile load. After one fibre breaks yarn bundles or untwisted cables tend to spread the stress load uniformly on the remaining unbroken fibres. This is the equal load share rule under which the load of the failed component is distributed equally among the remaining working components.

Coleman (1958) found the mean time to ultimate failure of a bundle of parallel fibres when the number of fibers becomes large. Birnbaum and Saunders (1958) derived the lifetime distributions of the materials. Phoenix (1978) showed that the system failure is asymptotically normally distributed as the number of components become large. A more general monotone load sharing rule assumes that the load on any individual component is nondecreasing as other items fail.

Apart from textile industry such model arises in manufacturing where a part can be considered failed only when the entire set of welded joints that holds the part together fails. However, the failure of one or two joints can increase the stress on remaining joints.

Kim and Kvam (2004) observed that such models also arise in sampling techniques. Suppose the total resourses allocated toward finding a finite set of items is fixed. Once one item is detected, resources can be redistributed to finding remaining items. This is a load sharing model. If the items are identical then an equal load share rule is the right one for studying system dependence.

Lynch (1999) characterised some relationships between the failure rate and the load share rule and Durham and Lynch (2000) studied relationships for some specified load-share rules.

Kim and Kvam (2004) consider a k component parallel system. Initially the components have identical distribution with constant failure rate θ. After the failure of the first component the modified failure rate of $k-1$ components changes to $\gamma_1\theta$, for some $\gamma_1 > 0$, and so on. They find maximum likelihood estimators of the k parameters $\theta, \gamma_1, \gamma_2, \ldots, \gamma_{k-1}$. They consider the estimation of parameters under monotone load sharing $1 \leq \gamma_1 \leq \gamma_2 \leq \ldots \leq \gamma_{k-1}$. They also derived a likelihood ratio test for testing equality of $\gamma's$ against the alternative that they are monotone.

Kvam and Pena (2005) discuss load sharing models which are special cases of dynamic models where performance of the system changes as components fail or their performance deteriorates. They consider a k-component parallel system under an equal load share model. When the first component fails the failure rate of remaining components changes from $r(t)$ to $\gamma_1 r(t)$ and so on. They find an estimator of the component baseline cumulative hazard function $R = -log(1-F)$ and discuss its asymptotic distribution.

McCool (2006) modelled the time to failure as a two parameter Weibull distribution. He proposed a test to test the hypothesis that the failure of the first component in a parallel system shortens the life of the remaining components of the same system. We look at a similar testing problem without making any assumption on the distribution of component life-times.

In all above examples failure of first component adversely affects the system performance. On the other side the detection of a bug in a software can help in the detection of other bugs. Once a critical fault in the software has been detected it can help in finding other bugs wich had earlier been undetected. Drummond et al (2000) carried out a study in a vertebrate species showing that selective deaths due to food shortage result in surviving offsprings receiving an increased share of an undiminished food supply. They observed littermates of the domestic rabbit Oryctolagus cuniculus and found that after individual pups died, the total daily milk weight obtained by the litter continued to be same. Hence the surviving pups consumed more milk and showed greater growth. This necessitates considering the other one sided test also.

We restrict ourselves to two component systems and study their life-lengths, identifiablity issues and bounds under various observations schemes in section 2. In section 3 we propose a conditional failure rate model which explicitly uses the concept of additional load on the surviving component after the failure of one. We also discuss the popular Gumbel (1960) and

Freund (1961) models from this point of view. Section 4 provides a test for the null hypothesis that the two components fail without any additional load on the surviving component against the alternative that there is such an additional load. In section 5 we extend the test to the case of right censored data. Section 6 includes a simulation study and in section 7 the tests are illustrated on real data. The last section has comments and conclusions.

2 Bounds under various observation schemes

In this section we look at various sampling schemes that arise with load sharing and discuss corresponding identifiability issues. If the component lifelengths are not identifiable, we propose bounds which can be estimated from the corresponding data.

2.1 Components are independent and identically distributed - observe both order statistics

Suppose that the component lifetimes U_1, U_2 are continuous, positive valued, independent random variables with a common distribution function $F(x)$, survival function $\bar{F}(x)$, density function $f(x)$, failure rate function $r_F(x) = \frac{f(x)}{\bar{F}(x)}$. Suppose that the lifetime of the component which fails first is given by $X = min(U_1, U_2)$ and the lifetime of the component which fails second, which is the same as the system lifelength, is $Y = max(U_1, U_2)$. Because U_1, U_2 are i.i.d random variables, it is easy to see that the joint density of the two ordered failure times X, Y is given by

$$\begin{aligned} g(x, y) &= 2\ f(x) f(y), \quad 0 < x < y < \infty, \\ &= 0, \quad otherwise. \end{aligned}$$

The marginal distribution of the minimum X and the maximum Y, respectively are given by

$$G(x) = 1 - (1 - F(x))^2, \ \ H(x) = [F(x)]^2,$$

and the density functions by

$$g(x) = 2f(x)(1 - F(x), \ \ h(x) = 2f(x)F(x).$$

Hence the marginal distribution of U_1 can be identified from the distribution of either of the order statistics .

2.2 Components are dependent - observe order statistics

Next suppose that the component lifetimes are no longer independent, as in the case of a pair of lungs or a pair of kidneys. Let the joint distribution function of component lifelengths (U_1, U_2) be given by $F(x, y)$ and joint pdf by $f(x, y)$. Its joint survival is given by

$$\bar{F}(x, y) = 1 - F_1(x) - F_2(y) + F(x, y). \tag{1}$$

As before $X = min(U_1, U_2)$ and $Y = max(U_1, U_2)$

Then, the joint distribution function of (X, Y) is given by

$$G(x, y) = F(x, y) + F(y, x) - F(min(x, y), min(x, y)), \quad 0 < x < y < \infty. \tag{2}$$

And the joint density function of (X, Y) is given by

$$g(x, y) = f(x, y) + f(y, x), \quad 0 < x < y < \infty. \tag{3}$$

We have identifiability only along the diagonal (x, x) .

Then the survival function of the first failure X is given by

$$\bar{G}(x) = P[X > x] = P[U_1, U_2 > x] = \bar{F}(x, x) = 1 - F_1(x) - F_2(x) + F(x, x). \tag{4}$$

The distribution function of the system lifelength Y is given by

$$H(x) = P[U_1, U_2 \leq x] = F(x, x). \tag{5}$$

Let $g(x)$ and $h(x)$ be the corresponding density functions.

When the component lifetimes are not independent, we cannot identify the joint distribution $F(x, y)$ from the joint distribution of the order statistics (X, Y). However, we have the following bounds.

Theorem 1:

$$G(min(x, y), min(x, y)) \leq F(x, y) \leq G(min(x, y), max(x, y)) \;\; \forall x, y \tag{6}$$

Proof : We can write

$$F(x, y) = G(x, y)P(X_1 \leq X_2) + G(y, x)P(X_2 \leq X_1).$$

Therefore

$$min\{G(x, y), G(y, x)\} \leq F(x, y) \leq max\{G(x, y), G(y, x)\}.$$

Separately considering $x < y$ and $y < x$ leads to the inequality given above.

Note that if $F(x, y)$ is exchangeable then $F(x, y) = F(y, x)$,

$$F(x, y) = \frac{1}{2}[G(x, y) + G(min(x, y), min(x, y)),$$

and

$$f(x, y) = \frac{1}{2}g(x, y).$$

Thus, when $F(x, y)$ is exchangeable, then it is identifiable on the basis of joint distribution $G(x, y)$, not otherwise. Independence of U_1, U_2 is neither sufficient nor necessary for identifiability.

2.3 Components are dependent - observe the maximum and its identifier

The third possible sampling scheme is the following. Suppose that the system life Y and the label of the component whose failure coincided with the system failure are observable. That is, we observe (Y, δ), where $\delta = 1$ if $Y = U_1$ and $\delta = 2$ if $Y = U_2$. Basu and Ghosh (1981) called it the complementary risks data and discussed identifiability issues associated with it. The joint distribution function of (Y, δ), is given by the following pair of sub-distribution functions $H(t, 1) = P(Y \leq t,\ \delta = 1)$ and $H(t, 2) = P(Y \leq t,\ \delta = 2)$. The joint distribution of (U_1, U_2) given by $F(x, y)$ is not identifiable under this sampling scheme. However we note that the following bounds hold.

$$H(min(x, y), 1) + H(min(x, y), 2) \leq F(x, y) \leq H(x, 1) + H(y, 2). \quad (7)$$

In case only the maximum Y is available and the markers given by δ are not known, then we have

$$H(min(x, y)) \leq F(x, y) \leq H(max(x, y)). \quad (8)$$

These results are analogous to those obtained by Peterson (1976) for series systems. Using the bounds in (6), (7) and (8) one can obtain conservative confidence bounds for $F(x, y)$ following the approach in Deshpande and Karia (1995). As suggested by them one needs to obtain the lower confidence limit (band) for the lower bound and upper confidence limit (band) for the upper bound. These limits (bands) can be based on consistent estimators of the respective bounds together with the asymptotic distributions of these estimators. The bivariate and the univariate empirical distribution function may be used as an estimators for $G(x, y)$ and $H(x)$, respectively and the empirical sub-distribution functions for $H(x, 1)$ and $H(y, 2)$.

3 A Model for load sharing

X, Y are the minimum and the maximum of U_1, U_2 which are i.i.d with distribution function $F(x)$. One can obtain the conditional density of Y

given $X = x$ as

$$h_{Y|X=x}(y) = f(y)/[1 - F(x)], \quad 0 < x < y < \infty.$$

Therefore, the joint density of X and Y can be equivalently defined by the pair consisting of the marginal density of X and conditional density of Y given X, that is, by

$$\{2f(x)(1 - F(x)),\ f(y)/[1 - F(x)]\ \}.$$

Then the failure rate of the conditional distribution of Y given X is

$$r_{Y|X=x}(y) = \frac{f(y)}{(1 - F(x))} \frac{1 - F(x)}{1 - F(y)} = \frac{f(y)}{1 - F(y)} = r_F(y), \quad 0 < x < y < \infty.$$

That is if we have a parallel system based on two independent and identically distributed components, then the failure rate of the system given the failure of the first component is the same as the failure rate of the original component. That is, the system failure rate is not affected by the failure of the first component.

Next consider the experimental situation where initially the components are independent and identically distributed but the first failure shifts the load to the surviving components. In such a case we expect that the conditional failure rate gets affected and we suggest a proportional hazards model as follows

$$r_{\Delta,Y|X=x}(y) = \Delta r_F(y), \quad 0 < x < y < \infty, \quad \Delta \geq 1. \tag{9}$$

$\Delta = 1$ gives the independence of component lives and the fact that the first failure shifts an extra load on the surviving component can be modeled by taking $\Delta > 1$.

Under this set up, the pair of marginal density functions of X and the conditional density of Y given $X = x$ is

$$\{2f(x)(1 - F(x)),\ \Delta f(y) \frac{(1 - F(y))^{\Delta - 1}}{(1 - F(x))^{\Delta}}\},\ 0 < x < y < \infty,\ 1 \leq \Delta < \infty,$$

and the joint density function of the ordered component lives is

$$g_{\Delta}(x,y) = 2\Delta f(x)f(y)\frac{(1-F(y))^{\Delta-1}}{(1-F(x))^{\Delta-1}},\ 0<x<y<\infty,\ 1\leq\Delta<\infty. \tag{10}$$

$$\begin{aligned} h_{\Delta}(y) \ &= \ \frac{2\Delta f(y)[\bar{F}(y)^{\Delta-1}-\bar{F}(y)]}{2-\Delta},\ \ \Delta\neq 2 \\ & \quad -4\bar{F}(y)log\bar{F}(y)f(y),\ \ \Delta=2,\ y>0. \end{aligned}$$

It is interesting to note that the unordered random variables corresponding to the joint density function $g_{\Delta}(x,y)$ are neither independent nor identically distributed.

As an illustration suppose that U_1, U_2 are i.i.d exponential random variables with failure rate λ. Then the joint density function of X, Y under the proposed model is given by

$$g_{\Delta}(x,y) = 2\Delta\lambda^2 exp-[\lambda(x+y)]exp-[\lambda(\Delta-1)(y-x)],\ 0<x<y<\infty. \tag{11}$$

The marginal density of the first component X and the system failure time Y is

$$g(x) = 2exp-[2x],\ x>0,$$

$$\begin{aligned} h(y) \ &= \ \frac{2\lambda\Delta[exp-[\lambda\Delta y]-exp-[2\lambda y]]}{2-\Delta},\ \ \Delta\neq 2 \\ & \quad 4\lambda^2 yexp-[2\lambda y],\ \ \Delta=2,\ \ y>0. \end{aligned}$$

The conditional density of Y given $X=x$ is

$$h_{Y|X=x}(y) = \Delta exp-[\Delta(y-x)],\ 0<x<y<\infty.$$

(10) gives us a new bivariate model which brings in the effect of load sharing. One could possibly obtain more models by starting with i.i.d Weibull and other distributions for the U_1, U_2. Further we may bring in the effect of loadsharing by a nonproportional conditional hazard rate as well. This opens up a rich class of bivariate models.

As mentioned earlier the original lifetimes U_1, U_2 could be dependent. We would like to compare the joint density of the minimum and maximum proposed in (10) with those arising in the case when the original lifetimes follow bivariate distribution due to Gumbel (1960) and Freund (1961). We look at the conditional distributions and the conditional failure rates.

The conditional density function of Y given $X = x$ is

$$h_{(Y|X=x)}(y) = \frac{g(x,y)}{g(x)} = \frac{f(x,y) + f(y,x)}{-d/dx(\bar{F}(x,x))}, \quad x < y. \tag{12}$$

Its conditional distribution function is

$$H_{(Y|X=x)}(y) = \frac{\int_x^y g(x,u)du}{g(x)}, \quad x < y. \tag{13}$$

The conditional survival function is

$$\bar{G}_{(Y|X=x)}(y) = 1 - \frac{\int_x^y g(x,u)du}{g(x)}, \quad x < y. \tag{14}$$

So that the conditional failure rate of the system given the first failure is

$$r_{Y|X}(y) = \frac{f(x,y) + f(y,x)}{g(x) - \int_x^y [f(x,u) + f(u,x)]du}, \quad x < y. \tag{15}$$

And the system failure rate is

$$r_Y(x) = \frac{d/dx(F(x,x))}{1 - F(x,x)}, \quad x > 0. \tag{16}$$

Now we work out these expressions for Gumbel's bivariate exponential given below.

$$F(x,y) = 1 - e^{-x} - e^{-y} + e^{-x-y-\delta xy}, \quad x, y > 0, \ 0 \le \delta \le 1.$$

Note that U_1 and U_2 are independent if $\delta = 0$. That essentially reduces to the case of i.i.d. exponentials random variables discussed earlier.

$$f(x,y) \ = \ [(1+\delta y)(1+\delta y) - \delta]e^{-x-y-\delta xy}, \quad x, y > 0, \ 0 \le \delta \le 1. \tag{17}$$

Hence the joint pdf of (X, Y) is given by

$$g(x, y) = 2f(x, y), \quad x < y.$$

Survival function and density function of the time to first failure X is

$$\begin{aligned} \bar{G}(x) &= e^{-2x-\delta x^2}, \\ g(x) &= (2 + 2\delta x)e^{-2x-\delta x^2}. \end{aligned} \tag{18}$$

Notice that the time to first failure has the linear failure rate.

And the distribution function and density function of the system failure Y is

$$H(x) = 1 - 2e^{-x} + e^{-2x-\delta x^2},$$

and

$$h(x) = 2e^{-x} - (2 + 2\delta x)e^{-2x-\delta x^2}, \; x > 0.$$

The conditional density and distribution of Y given $X = x$ are given as follows

$$h_{Y|X=x}(y) = \frac{[(1+\delta y)(1+\delta y) - \delta]e^{-x-y-\delta xy}]}{(1+\delta x)e^{-2x-\delta x^2}}, \; x < y,$$

$$H_{Y|X=x}(y) = \frac{e^{-x-y-\delta xy} - e^{-2x-\delta x^2} - \delta y e^{-x-y-\delta xy} + \delta x e^{-2x-\delta x^2}}{2(1+\delta x)e^{-2x-\delta x^2}}, \; x < y.$$

And the conditional failure rate of Y given $X = x$ for $x < y$ is

$$r_{Y|X=x}(y) = \frac{[(1+\delta y)(1+\delta y) - \delta]e^{-x-y-\delta xy}]}{2(1+\delta x)e^{-2x-\delta x^2} - e^{-x-y-\delta xy} + e^{-2x-\delta x^2} + \delta y e^{-x-y-\delta xy} - \delta x e^{-2x-\delta x^2}}.$$

Finally the failure rate of Y is

$$r_Y(x) = \frac{2e^{-x} - (2 + 2\delta x)e^{-2x-\delta x^2}}{2e^{-x} - e^{-2x-\delta x^2}}, \quad x > 0.$$

Note that U_1, U_2 have failure rate 1.

As another example we consider Freund's distribution. The motivation of the model was to start with independent exponentials and consider the modifications in case there is load sharing. Note that the authors were not looking at the joint distribution of the minimum and the maximum which is of interest to us. The joint density function of U_1, U_2 is given by

$$f(x,y) = \begin{cases} \alpha\beta' exp[-\beta' y - (\alpha+\beta-\beta')x] & 0 < x < y \\ \beta\alpha' exp[-\alpha' x - (\alpha+\beta-\alpha')y] & 0 < y < x. \end{cases}$$

Then, the distribution of the system failure X is given by

$$\begin{aligned} F(x,x) &= \frac{\alpha}{\alpha+\beta-\beta'}(1-exp[-\beta' x]) + \frac{\beta}{\alpha+\beta-\alpha'}(1-exp[-\alpha' x]) \\ &- \frac{\alpha\beta'}{(\alpha+\beta)(\alpha+\beta-\beta')}(1-exp[-(\alpha+\beta)x]) \\ &- \frac{\beta\alpha'}{(\alpha+\beta)(\alpha+\beta-\alpha')}(1-exp[-(\alpha+\beta)x]). \end{aligned}$$

The density of the maximum is

$$g_Y(x) = \frac{\alpha\beta'}{\alpha+\beta-\beta'} exp[-\beta' x] + \frac{\beta\alpha'}{\alpha+\beta-\alpha'} exp[-\alpha' x], \quad x > 0.$$

And the density of the minimum is

$$g_x(x) = (\alpha+\beta) exp[-(\alpha+\beta)x], \quad x > 0.$$

Hence the minimum is exponential with failure rate $\alpha+\beta$.

The joint density of the two order staistics (X, Y) is given by

$$g(x,y) = \alpha\beta' exp[-\beta' y - (\alpha+\beta-\beta')x] + \beta\alpha' exp[-\alpha' y - (\alpha+\beta-\alpha')x], \quad 0 < x < y.$$

After simple but lengthy calculations, the conditional failure rate of the maximum, given $X = x$ is

$$r_{Y|X=x}(y) = \frac{\alpha\beta' exp[-\beta'(y-x)] + \beta\alpha' exp - [\alpha'(y-x)]}{\alpha exp[-\beta'(y-x)] + \beta exp[-\alpha'(y-x)]}, \quad x < y.$$

Since we are looking at the case when U_1, U_2 are identically distributed but are not independent, we take $\alpha = \beta, \alpha' = \beta'$. It is interesting to note that under this assumption, $r_{Y|X=x}(y) = \alpha'$. This is not equal to the failure rate of any component of the system.

Hence if the component lifelengths are identical, the conditional failure rate of the system given the first failure is a constant. That is, the conditional distribution of $Y|X = x$ is exponential with failure rate α'. Observe that in this case the distribution of the minimum is exponential with failure rate 2α.

In none of these cases can we write $r_{\Delta,Y|X}(y) = \Delta r_F(y), \quad 0 < x < y < \infty, \quad \Delta \geq 1$. Hence the joint density function of the order statistics given by (10) appears to be a more meaningful way of looking at load sharing instead of the joint distribution of order staistics arising out of Gumbel and bivariate distributions since it takes care of the load sharing mechanism and proposes a way to deal with it.

It is interesting to observe the following special case. Suppose U_I, U_2 are independent but not identically distributed with density functions $f_1(x)$ and $f_2(x)$, respectively.

Then joint density of of minimum and maximum is

$$g(x, y) = f_1(x)f_2(y) + f_1(y)f_2(x),\ x < y. \tag{19}$$

The density of the minimum X is

$$g(x) = f_1(x)\bar{F}_2(x) + \bar{F}_1(x)f_2(x),\ x > 0. \tag{20}$$

Hence conditional failure rate of maximum given minimum is

$$r_{Y|X=x}(y) = \frac{f_1(x)f_2(y) + f_1(y)f_2(x)}{f_1(x)\bar{F}_2(y) + \bar{F}_1(y)f_2(x)},\ x < y. \tag{21}$$

Suppose U_1 has exponential distribution with failure rate 1 and U_2 has exponential with failure rate $\lambda > 1$. Then conditional failure rate of maximum given minimum is

$$r_{Y|X=x}(y) = \frac{\lambda e^{-x}e^{-\lambda y} + \lambda e^{-\lambda x}e^{-y}}{e^{-x}e^{-\lambda y} + \lambda e^{-\lambda x}e^{-y}},\ \ x < y. \tag{22}$$

Then for $\lambda > 1$ we have

$$1 \leq r_{Y|X=x}(y) \leq \lambda. \tag{23}$$

In general one can show that if $r_{U_1}(y) \leq r_{U_2}(y)$, then

$$r_{U_1}(y) = \frac{f_1(y)}{\bar{F}_1(y)} \leq r_{Y|X=x}(y) \leq \frac{f_2(y)}{\bar{F}_2(y)} = r_{U_2}(y). \tag{24}$$

4 A Test

Our next question is whether the first failure affects the failure of the system or not. Hence we wish to test the null hypothesis H_0 that the first failure does not affect the system lifelength against the alternative H_1 that the first failure stochastically reduces the system lifelength. Or equivalently, one may say that under H_0 first and second failure times are from the order statistics distribution based on i.i.d random variables and under H_1 the second failure occurs earlier than what is predicted by the order statistics distribution based on i.i.d random variables. Thus under the alternative the distribution function of system life Y given $X = x$ is decreasing in x.

Suppose the data consists of n independent pairs of ordered component lifetimes (X_i, Y_i), $i = 1, 2, ..., n$. Trivially, $X_i < Y_i$, $i = 1, 2, \ldots, n$. Consider the following U-statistics

$$U(X_1, Y_1, \ldots, X_n, Y_n) = \frac{1}{\binom{n}{2}} \sum_{i<j} \frac{1}{2}[h(X_i, Y_j) + h(X_j, Y_i)]. \tag{25}$$

where $h(X_i, Y_j) = I(X_i < Y_j)$. Then, it is easy to see that $E(U) = \int_0^\infty G(x)dH(x)$. Under H_0 it is equal to $\frac{5}{6}$ and under the model (10), it is less than $\frac{5}{6}$ for $\Delta > 1$, leading to consistency of the tests. Under H_0, we have

$$Var(U) = \frac{1}{9n(n-1)} + \frac{4(n-2)}{90\ n(n-1).}$$

¿From the limiting theorem of U-statistics (Serfling (1980)) it follows that

$$\frac{\sqrt(n)(U - 5/6)}{\sqrt{2/45}} \xrightarrow{d} N(0, 1).$$

Small values of the statistic are significant. It is interesting to note that under the set up when component lifetimes are independent, the null mean and variance are distribution free.

Let R_j be the rank of Y_j in the combined arrangement of all the minimum and maximum. Then, we can express the U-statistics as function of the ranks as follows,

$$\binom{n}{2}U = \frac{1}{2}\sum_{j=1}^{n} R_j - \frac{n(n+3)}{4}. \tag{26}$$

5 The Censored Case

In almost all survival studies complete data is not available due to presence of a censoring mechanism. However the censoring can occur in several ways. The monitoring starts after the first failure has already occured (that is, one kidney has already failed). Here the minimum X is left censored by the age at which the monitoring begins and the maximum is observed without censoring. It is also possible that the first failure X is observed but the maximum is right censored by death due to other causes. In the extreme case it is possible that the minimum X is left censored and the maximum Y is right censored. We will only look at the right censoring case.

5.1 Right Censoring

Suppose that U_1 and U_2 are independent and identically distributed random variables. Let C, with distribution function $K_R(t)$, denote the random variable which censors Y. It acts independent of the pair (X, Y). Based on a random sample from this set up we consider a kernel

$$h_C(X_i, Y_i, X_j, Y_j) = \frac{1}{2}[I(X_i < min(Y_j, C_j)) + I(X_j < min(Y_i, C_i))]. \tag{27}$$

Let U_{CR} be the U-statistic estimator of the kernel $h_C(X_i, Y_i, X_j, Y_j)$.

Assume that the distribution of C_R satisfies the Koziol-Green model with $K_R(x) = [F(x)]^\theta$ $\theta \neq 2$.

Note that $\theta = 2$ gives the distribution function of Y, for $\theta > 2$, we have $K_R(x) > H(x)$, that is, the censoring random variable is stochastically smaller than Y and this indicates heavy censoring. The reverse is true if $\theta < 2$.

Under this set up we have

$$\begin{aligned} E[U_C] &= \int_0^\infty (1-F^2(x))(1-F^\theta(x))2\bar{F}(x)f(x)dx \\ &= 2[\frac{5}{12} - \frac{1}{\theta+1} + \frac{1}{\theta+2} + \frac{1}{\theta+3} - \frac{1}{\theta+4}]. \end{aligned} \tag{28}$$

$E(U_C)$ depends on unknown θ. However, note that $P[Y < C] = \frac{\theta}{\theta+2}$. One can easily replace the unknown θ by its consistent estimator $\frac{2A}{1-A}$, where A is the proportion of Y observations that are uncensored. Let us call this $\hat{E}(U_C]$.

The asymptotic variance of U_C is a lengthy (there are around 50 terms) expression involving θ and hence is not being reported. It can be estimated consistently by replacing θ by its consisitent estimator.

We can also estimate the asymptotic variance σ_C^2 as follows.

$$\begin{aligned} \sigma_C^2 &= \frac{1}{4}[E[\bar{H}(X)\bar{K}(X)]^2 + E[G(min(Y,C))]^2 \\ &\quad +2E[(\bar{H}(X)\bar{K}(X))G(min(Y,C))]] - E^2[U_C]. \end{aligned} \tag{29}$$

Then

$$\begin{aligned} E[\bar{H}(X)\bar{K}(X)]^2 &= \int \bar{H}^2(x)\bar{K}^2(x)dG(x) \\ &= P(Min(Y_2,C_2) > X_1, Min(Y_3,C_3) > X_1) \ \ (A_1 \ \ say), \end{aligned} \tag{30}$$

$$\begin{aligned} E[G(min(Y,C))]^2 &= \int G^2(min(y,c))dH(y)dK(c) \\ &= P[X_2 < min(Y_1,C_1), X_3 < min(Y_1,C_1)] \ \ (A_2 \ \ say), \end{aligned} \tag{31}$$

$$\begin{aligned} & E[\bar{H}(X)\bar{K}(X)G(min(Y,C))] \\ = & \int \bar{H}(x)\bar{K}(x)G(min(y,c))g(x,y)k(c)dxdydc \\ = & P[Min(Y_2,C_2) > X_1, X_3 < min(Y_1,C_1), Y_1 > X_1] \; (A_3 \; say). \end{aligned} \tag{32}$$

Probability expressions given in A_1, A_2, A_3 can be estimated unbiasedly by respective indicator functions. One can look at symmetric versions of these indicator functions and define corresponding U-statistics. Thus, we have U-statistics estimator for the asymptotic variance. Using the results of U-statistics we have $\frac{\sqrt{n}(U_C - \hat{E}(U_C]}{2\hat{\sigma}_C}$ has limiting normal distribution.

6 Simulations

We carried out a simulation study to look at the power of the test and also the level attained. First we draw random samples of X, Y with sample size $n = 50, 100, 200$ from the joint pdf given by (10). Let $F(x) = 1 - e^{-x}$. The procedure was repeated 1000 times. Table 1 gives the power. The values below the line look at $\Delta < 1$ and hence the rejection was for large values. This essentially takes care of the case when the system performance improves after first failure.

¿From Table 1 we see that the distribution is slightly skew for small sample values since the expectation of the statistic is 5/6 under the null hypothesis, while the range is [0,1]. This results in slow convergence to .05, the asymptotic level for the case $\Delta = 1$. Otherwise the power inreases with increase in sample size and as Δ moves away from 1.

Next we drew random samples from i.i.d exponentials with sample size $n = 50, 100, 200$. Under the alternative the components continue to be independent but not identically distributed. In this case if we are looking for departures from the i.i.d. structure, $E(U) > 5/6$. Hence we reject for large values. Table 2 reports the power when the experiment is repeated 1000 times.

When $\lambda_1 = \lambda_2$, both components are independent and identically distributed and hence components are i.i.d. Otherwise they are independent but not identically distributed. For all sample sizes the exact level attained is close to .05. Unequal vaues of λ denote departure from i.i.d. set up. The higher the difference between the two values of λ, the higher is the power. Similarly, the power increases with the increase in sample size.

Finally, we generate random samples from bivariate Gumbel distribution (1960). The joint survival distribution of the components is given by

$$\bar{F}(x, y) = exp[-\lambda_1 x - \lambda_2 y - \lambda_3 xy], \; x, y > 0, \tag{33}$$

$\lambda_1, \lambda_2 > 0, 0 < \lambda_3 < \lambda_1\lambda_2$. When $\lambda_3 = 0$, the components are independent and and have exponenential marginals. When $\lambda_1 = \lambda_2$, the two components are identically distributed . Hence the case $\lambda_1 = \lambda_2, \lambda_3 = 0$ correspond to the i.i.d set up and all other values indicate departure from the i.i.d set up. Here again the rejection is for large values of the statistic.

Table 1: Power when sample from joint distribution of X, Y

n	Δ	power
50	1.0	0.061
100	1.0	0.052
200	1.0	0.044
50	1.10	0.098
100	1.10	0.119
200	1.10	0.192
50	1.20	0.191
100	1.20	0.251
200	1.20	0.395
50	1.30	0.267
100	1.30	0.412
200	1.30	0.619
50	1.40	0.391
100	1.40	0.573
200	1.40	0.830
50	1.50	0.466
100	1.50	0.733
200	1.50	0.929
50	0.90	0.118
100	0.90	0.124
200	0.90	0.184
50	0.80	0.193
200	0.80	0.507
50	0.70	0.348
100	0.70	0.562
200	0.70	0.830
100	0.60	0.821
200	0.60	0.967
50	0.50	0.768
100	0.50	0.954
200	0.50	1.000

Table 2: Power when sample from independent exponentials

n	λ_1	λ_2	power
50	1	1	0.049
100	1	1	0.044
200	1	1	0.049
50	1	2	0.133
100	1	2	0.182
200	1	2	0.311
50	1	3	0.359
100	1	3	0.583
200	1	3	0.860
50	1	4	0.626
100	1	4	0.875
200	1	4	0.990
50	2	4	0.133
100	2	4	0.182
200	2	4	0.311
50	4	2	0.140
100	4	2	0.191
200	4	2	0.306
50	3	1	0.349
100	3	1	0.721
100	3	1	0.830

Table 3: Power when sample from bivariate Gumbel

n	λ_1	λ_2	λ_3	power
50	1	1	0	0.049
100	1	1	0	0.055
200	1	1	0	0.048
50	1	2	0	0.146
100	1	2	0	0.197
200	1	2	0	0.305
50	1	1	0 1	0.104
100	1	1	0.1	0.150
200	1	1	0.1	0.217
50	1	2	0.1	0.189
100	1	2	0.1	0.325
200	1	2	0.1	0.503
50	1	1	0.2	0.209
100	1	1	0.2	0.303
200	1	1	0.2	0.502
50	1	2	0.2	0.251
100	1	2	0.2	0.437
200	1	2	0.2	0.720

7 Examples

The following situations are examples of load sharing in biological and other disciplines.

Example 1: Mantel, Bohidar and Ciminera (1977) report data on 50 male and 50 female litters, each of three rats. One rat in each litter was drug-treated and the other two served as control animals. The records are either on the week of tumor appearance or the week of death. To test whether the death of a littermate affects the lifetime of the surviving animal, we use the test in Section 4 for the two control animals. We delete the litters with tumor deaths and also the litters in which both the control animals were sacrificed or died at the end of 104 weeks, the time at which the study was ended. Thus we have 48 male litters and 22 female litters. The values of the test statistic for the male litters is 0.044, and for the female litters is 2.504. In case of the male litters, the null hypothesis H_0 that the first death does not affect the second death is not rejected , whereas in case of the female litters, it is rejected even at 1% level of significance in favor of H_1' : death of a littermate increases the residual life of the surviving mate.

The above example indicates that in litter-matched case control sudies, the lifetimes of the littermates may have to be treated as bivariate data as in Hougaard (2000).

Example 2: We consider data from an article of Kvam and Peña (2003) on three star players in a basketball team. The data are from the Basketball Association franchise Boston Celtics obtined during the second half of the 2001-2002 season. The data, given in Table 4, consist of the game times for each player's second personal foul for the games in the season in which all three players started the game and committed at least two fouls by the end. Kvam and Peña conjecture that once a player commits two fouls (and is likely to be out of the game for a period of time), the foul rate of the other star players will change. Either the foul rate might decrease if all the players decide to play conservatively or might increase if the other star players have to shoulder the responsibility on defense and thus are more prone to foul. They consider that the three star players compose a system and define a component failure as the event when a player commits two fouls. Thus the failure time of a player is same as the time-until-second-foul for the player.

Here we consider that the system comprises of two star players and obtain the value of the test statistic of Section 4 for the three possible combinations.

The value of the test statistic for players I and II is -0.332 and for players II and III, the value is 0.299. Thus in both these cases there is no significant evidence against H_0 at 1% level of significance. Whereas for players I and III, the value of the test statistic is -1.461 , which is significant only at 10% level in favour of H_1. We may conclude that two fouls of players I and III affect each other moderately and after one commits two fouls the other player is quite likely to commit his second foul sooner than he would have committed otherwise.

Example 3: For the ball bearings data given in McCool (2006), the value of the test statistic is -2.324. Thus we reject H_0 in favour of H_1 : failure of one ball bearing increases the load on the other ball bearings. However the data are available for six systems only.

Table 4: Time until second personal foul in 28 basketball games

Game	Player I	Player II	Player III	Game	Player I	Player II	Player III
1	21.02	30.22	43.43	15	42.06	23.21	45.36
2	24.25	45.54	17.19	16	28.51	33.59	16.2
3	6.555	19.47	23.28	17	34.56	32.53	40.44
4	15.35	16.37	25.4	18	40.33	15.35	28.33
5	39.08	30.32	43.53	19	27.56	46.21	28.05
6	16.2	4.16	39.52	20	9.54	36.21	28.12
7	34.59	46.44	16.33	21	27.09	11.11	23.33
8	19.1	38.4	20.17	22	40.36	33.21	17.04
9	28.22	37.43	25.41	23	41.44	36.28	19.13
10	32	45.52	39.11	24	32.23	8.17	41.27
11	11.25	19.09	11.59	25	7.53	37.31	13.43
12	17.39	25.43	22.51	26	28.34	35.58	41.48
13	28.47	31.15	2.41	27	26.32	28.02	29.33
14	23.42	31.28	40.03	28	30.47	40.4	42.13

8 Conclusions

Model proposed in (10) incorporates the changes in the performance of a two component system due to the failure of the first component. These models, for various choices of F, give us families of bivariate distributions which incorporate load sharing ideas better than the existing bivariate models as those due to Gumbel and Freund. One could also look at nonproportional models. We are looking at nonparametric estimation of the hazard rates of the proposed model in the presence of covariates.

On failure of one component, the surviving component may either have to undergo extra stress, leading to stochastically shorter residual lifetimes, or have access to extra resources, leading to longer residual lifetimes than what is warranted under independence. The extra load on the surviving components is observed in most mechanical systems and also in organic systems such as the two kidneys. On the other hand if two foraging animals have access to a fixed stock of food, then the death of one will make all the remaining food available to the surviving animal therby reducing the load.

Besides one could look at testing problem in context of a k component system where the failure of a subset of size k_1 (say) random units affects the lifetime of remaining $(k - k_1)$ units. Then there is the possibility of constructing Kolmogrov-Smirnov type tests in these situations. All these problems are being considered and further work will be reported when completed.

References

Basu, A. P. and Ghosh, J. K. (1980). Identifiability of distributions under competing risks and complementary risks model. *Comm. Statist. A-Theory Methods* **9** , 1515–1525.

Birnbaum, Z. W. and Saunders, S. C. (1958) A statistical model for life-length of materials. *J. Amer. Statist. Assoc.* **53**, 151-160.

Coleman, B.D. (1958). Statistics and time dependence of mathematical breakdowns in fibres. *J. Applied Physics* **29**, 968-983.

Daniels, H. E. (1945). The statistical theory of the strength of bundles of threads. *I. Proc. R. Soc. Lond. A* **183**, 404-435.

Deshpande J. V. and Karia S. (1995). Confidence bounds for the joint survival function in the dependent competing risks setup. *Commun. Statist. - Theory Meth.* **24**, 2315-2327.

Drummond, H., Vazquez, E., Sanchez-Colon, S., Martinez-Gomez, M. and Hudson, R. (2000). Competition for milk in the domestic rabbit: survivors benefit from littermate deaths. *Ethology* **106**, 511-526.

Durham, S. D.and Lynch, J. D.(2000). A threshold representation for the strength distribution of a complex load sharing system. *J. Statist. Plann. Inference* **83** , 25-46.

Freund, J. E. (1961). A bivariate extension of the exponential distribution. *J. Am. Statist. Assoc.* **56**, 971-977.

Gross, A.J., Clark, V.A. and Liu, V. (1971). Estimation of survival parameters when one of two organs must function for survival. *Biometrics* **27**, 369-377.

Gumbel, J. E. (1960). Bivariate exponential distributions. *J. Amer. Statist. Assoc.* **55**, 698-707.

Hougaard, P. (2000). *Analysis of Multivariate Survival Data.* Springer. Kim, H. and Kvam, P. H. (2004). Reliability estimation based on system data with an unknown load share rule. Lifetime Data Anal. **10** , 83-94.

Kvam, P.H. and Pena, E.A. (2003). Estimating load sharing properties in a dynamic reliability system. *www.stat.sc.edu/ pena/TechReports/KvamPena2003.pdf.*

Kvam, P.H. and Pena, E.A. (2005). Estimating load sharing properties in a dynamic reliability system. *J. Amer. Statist. Assoc.* **100**, 263-272.

Lynch, J. D.(1999) . On the joint distribution of component failures for monotone load-sharing systems. *J. Statist. Plann. Inference 78* **78**, 13-21.

Mantel N., Bohidar, N. R. and Ciminera J. L. (1977). Mantel-Haenszel analysesf litter-matched time to response data, with modifications for recovery of interlitter information. *Cancer Res.* **37**, 3863-3868.

McCool, J. I. (2006). Testing for dependency of failure times in life testing. *Technometrics* **48**, 41-48.

Peterson, A. V. (1976). Bounds for the joint distribution function with fixed sub-distribution functions:Application to competing risks. *Proceedings of the National Academy of Science,* USA **73**, 11-13.

Phoenix, S. L. (1978). The asymptotic time to failure of a mechanical system of parallel members. *Siam J. Appl. Math.* **34**, 227-246.

Rosen, B.W. (1964). Tensile failure of fibrous composites. *AIAA J.* **2**, 1985-1991.

Serfling, R. J. (1980). *Approximation theorems of mathematical statistics.* John Wiley, New York.