Analysis of Censored Data
IMS Lecture Notes – Monograph Series (1995) Volume 27

The marginal distributions of lifetime variables which right censor each other

Tim Bedford
Department of Technical Mathematics and Informatics
Delft University of Technology
Mekelweg 4, 2628 CD Delft, The Netherlands

Isaac Meilijson
School of Mathematical Sciences
Raymond and Beverly Sackler Faculty of Exact Sciences
Tel Aviv University
Tel Aviv 69978, Israel

Abstract

The failure time of a machine with two modes of failure can be modeled as the minimum of two failure times (each associated with one failure mode). However, this model is not unique, i.e., as shown by Tsiatis, the joint distribution of the failure time and failure mode of the machine does not characterise the joint distribution - nor the marginal distributions - of the failure times under the two competing risks. Peterson introduced pointwise sharp bounds for these marginals. Crowder recognized that these bounds are not functionally sharp and restricted the class of functions containing all feasible marginals. In another publication, the authors improved on these bounds via a functional characterisation of the set of feasible pairs of marginals. As it turns out, not only is each marginal distribution function bounded below by its corresponding lower bound, but also its density is bounded below by the derivative of this bound. We present a summary of these results, describe a statistical application to the construction of confidence bands, and concentrate on the analysis of its rate of consistency under specific examples.

1 Introduction

Let T be the failure time of a machine with two modes of failure and let I indicate its mode of failure, i. e., $I = \{1\}, \{2\}$ or $\{1,2\}$ depending on whether failure of the machine was due to mode $1, 2$ or both simultaneously.

It is customary in Competing Risks to model the observed data (T, I) in terms of partially observed existing or conceptual variables X_1 and X_2 as $T = \min(X_1, X_2), I = \{i | X_i = T\}$, and assign to X_1 and X_2 the role of failure times associated to the failure modes.

It is well known that

- For every joint distribution $\mathcal{L}$ of (T, I) there exist joint distributions of (X_1, X_2) for which $(\min(X_1, X_2), \{i | X_i = \min(X_1, X_2)\})$ is distributed $\mathcal{L}$, and in fact (see Tsiatis [11]) there are many such joint distributions.
- For every joint distribution $\mathcal{L}$ of (T, I) with non-atomic T and $P(I = \{1, 2\}) = 0$, there exist (see Kaplan and Meier [4], Nadas [8], Miller [7]) *independent* random variables X_1 and X_2 for which $(\min(X_1, X_2), \{i | X_i = \min(X_1, X_2)\})$ is distributed $\mathcal{L}$. If $P(T \leq t, I = \{1\})$ and $P(T \leq t, I = \{2\})$ are simultaneously (in t) less than 1 or equal to 1, the distributions of these independent X_1 and X_2 are uniquely determined by $\mathcal{L}$. Conditions on $\mathcal{L}$ with positive $P(I = \{1, 2\})$ under which such independent X_1 and X_2 exist are also known (see Langberg, Proschan and Quinzi [5]).

The assumption of independence can be reasonable under certain circumstances. Consider for example a component which in order to function requires a continuous supply of electricity from the public network. Failure arising from a breakdown of the network may be reasonably considered independent of "intrinsic" failure of the component. However, it would be a poor engineer who preventively maintains independently of the component lifetime.

Accepting non-identifiability as a fact of life(time), Peterson [10] presents bounds for the joint distribution of X_1 and X_2 as well as for their marginals, assuming that $P(T = \{1, 2\}) = 0$. Peterson's bounds are sharp under continuity of the two sub-survival functions, in the sense that through every point between the two corresponding functions there passes the marginal distribution function of the corresponding lifetime variable, for some feasible joint distribution. Crowder [3] has found that these *pointwise* sharp bounds are not *functionally* sharp, that is, not every distribution function between the Peterson bounds is the marginal distribution of some feasible joint distribution: a feasible marginal leaves a nondecreasing gap with its corresponding lower Peterson bound. In attempting to add further conditions to obtain a characterisation, Crowder gives up on necessity by requiring a technically convenient but unnecessary condition, and unfortunately rules out sufficiency as well by failing to notice a pathological aspect of the upper Peterson bound: marginal distributions may only touch it lightly, in a way made precise by Bedford and Meilijson [1], who present a characterisation of the class of feasible marginal distributions. This results of Peterson, Crowder and Bedford & Meilijson are the basis of a new statistical test which can be used to test hypotheses about or build confidence bands for the marginal distributions.

In this paper we summarise the presentation of the functional bounds and the application of the statistical test, and develop some examples with the goal of understanding the degree of consistency of these tests. We will see that convergence can be extremely slow when X_2 is of a "quality control" nature with respect to X_1: $X_2 = 0$ if $X_1 \leq x_0$ and $X_2 = \infty$ otherwise, that is X_1 is observed if and only if $X_1 > x_0$. If X_1 is postulated to be marginally exponentially distributed, we obtain a confidence interval for its failure rate. This failure rate is uniquely identifiable if and only if the mean of X_1 (inverse failure rate) that generated the data is at least x_0. Letting n be the sample size of failed machines, if the mean of X_1 strictly exceeds x_0, the left endpoint of the confidence interval is $n^{-\frac{1}{2}}$ - consistent, whereas in the borderline identifiable case where the mean of X_1 is exactly equal to x_0, the left endpoint closes in at the slow rate $n^{-\frac{1}{6}}$.

The characterisation of the boundary of identifiability in the quality control example will be generalised from exponential distributions to families of distributions ordered by monotone likelihood ratio.

In many reliability textbook discussions of right censoring, after a brief warning that an independence assumption has been made, the Kaplan-Meier estimator is given as "the" way of estimating the required marginal. As a word of warning against the excessive use of the independence assumption, we will show that it may provide *extremely optimistic* assessments of lifetime distributions.

2 The bounds

Definition 1 *Let X_1 and X_2 be two random variables. The* distribution function *of X_i is $F_i(t) = P(X_i \leq t)$. The* sub-distribution function *of X_1 is the function $G_1(t) = P(X_1 \leq t, X_1 < X_2)$. Similarly we define $G_2(t) = P(X_2 \leq t, X_2 < X_1)$ and $G_{12}(t) = P(X_1 \leq t, X_2 = X_1)$.*

Let

$$\begin{aligned} \underline{F}_1(x) &= G_1(x) + G_{12}(x) = P(X_1 \leq x, X_1 \leq X_2) \\ \underline{F}_2(x) &= G_2(x) + G_{12}(x) = P(X_2 \leq x, X_2 \leq X_1) \end{aligned} \tag{1}$$

and define

$$\begin{aligned} \overline{F}_1(x) &= \underline{F}_1(x) + G_2(x) \\ \overline{F}_2(x) &= \underline{F}_2(x) + G_1(x). \end{aligned} \tag{2}$$

The *Peterson bounds* for the marginal distributions are

$$\underline{F}_i(x) \leq F_i(x) \leq \overline{F}_i(x), \quad \text{for all } x \geq 0 \text{ and } i = 1, 2 \ . \tag{3}$$

Since the sub-distribution functions G_i $(i = 1, 2)$ and G_{12} can be estimated from the observable data, so can the functions $\underline{F}_i$ and $\overline{F}_i$.

The bounds in (2) can be slightly improved in the presence of atoms by defining $\overline{F}_1(x)$ (respectively, $\overline{F}_2(x)$) to be $\overline{F}_1(x) = \underline{F}_1(x) + G_2(x-)$.

These bounds are *pointwise* bounds. It is obvious that not all functions K satisfying $\underline{F}_1(x) \leq K(x) \leq \overline{F}_1(x)$ for all x are distribution functions. It is also quite easy to see that not all distribution functions satisfying this inequality are allowable marginals. Indeed, if we consider the gap between each marginal distribution function and its lower bound,

$$\begin{aligned} F_1(x) - \underline{F}_1(x) &= P(X_1 \leq x, X_2 < X_1) \\ F_2(x) - \underline{F}_2(x) &= P(X_2 \leq x, X_1 < X_2) \end{aligned} \tag{4}$$

then, as observed by Crowder [3], these gaps must be nonnegative *and non-decreasing* as a function of x. In other words, feasible marginal distributions F_i are *co-monotone* with their lower Peterson bound $\underline{F}_i$.

The main result of Crowder [3], in the formulation corrected by Bedford and Meilijson [1], asserts that any distribution functions F_1 and F_2 greater or equal than, and co-monotone with, their respective Peterson lower bounds (3), and strictly less than their respective Peterson upper bounds (3), are marginals of the joint distribution of a pair (X_1, X_2) having the sub-distribution functions G_1, G_2 and G_{12}. The authors also characterise the extent to which the functions F_i can satisfy the Peterson upper bounds *with equality* and still qualify as feasible marginals, a question that may be safely ignored for most practical applications. This question will be touched upon in Section 4.

We illustrate this result with an example. Suppose we have failure data as follows:

Observed failure times, X_1	2, 4.5, 5, 7.5, 7.8
Censored times, X_2	1, 1.5, 3.1, 4.2, 6.1.

The data may be used to produce (empirical) sub-distribution functions as shown in Figure 1. Also shown on the figure is the empirical distribution function of $\min(X_1, X_2)$. We want to know what the "true" empirical distribution function of X_1 would have been if we had been able to observe it. Clearly the only thing we can say with any certainty is that each unobserved failure occurred *after* the corresponding censoring time. This implies that the (unseen) empirical distribution function

1. lies between the sub-distribution function for X_1 and the distribution function of the minimum, and

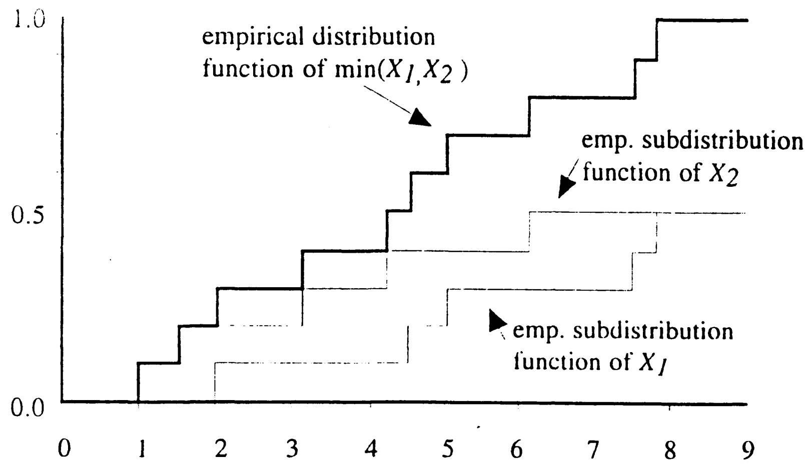

Figure 1: Empirical sub-distribution functions

2. has a jump wherever the sub-distribution function for X_1 has a jump.

This gives a wide range of possibilities for the marginal of X_1. In Figure 2 we show the sub-distribution function for X_1, the distribution function of $\min(X_1, X_2)$, and two possible distribution functions for X_1 (one with high correlation between X_1 and the censoring variable X_2, and one which assumes independence between X_1 and X_2).

We assert that any empirical distribution function $\hat{F}_1$ satisfying the conditions above *could* be the "true" unseen empirical distribution function of X_1 (and a corresponding statement holds also for X_2). To show this we have to construct a joint distribution between X_1 and X_2 displaying this feasibility. This is done in Bedford and Meilijson [1].

3 A statistical test

Given that non-identifiability is a fact of life, and that the assumption of independence between preventive maintenance time and failure time seems (even) less realistic than an assumption that the underlying failure time is exponentially distributed, we may ask the question "Which exponential lifetime parameters (constant failure rates) may be excluded given the data?"

A statistical test has been developed based on the Kolmogorov Smirnov statistic (see [1]) which enables us to exclude certain hypothesised distributions from being the marginal distribution of X_1. The test is distribution-free and can therefore be applied to any hypothesised distribution, not just to those from the exponential family.

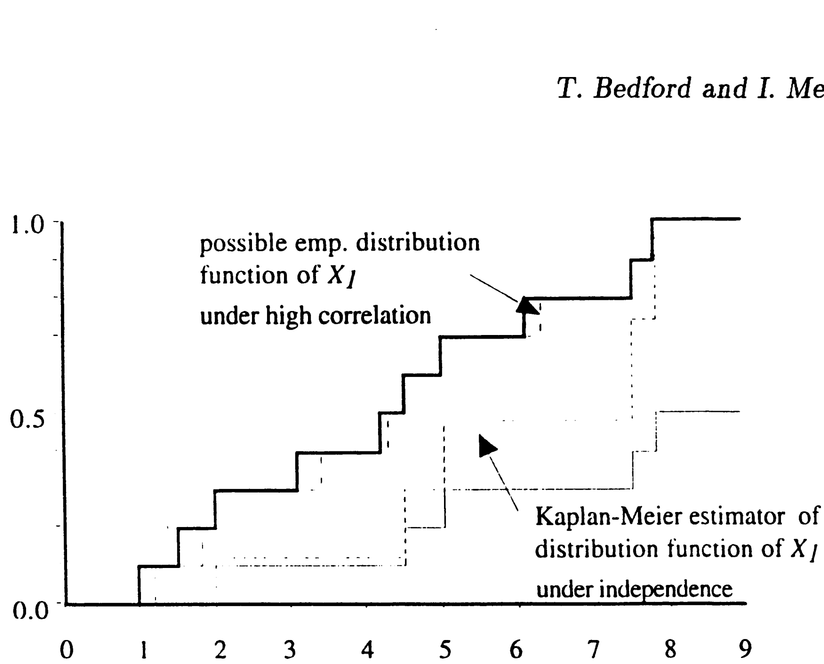

Figure 2: Two possible marginal distribution functions

The Kolmogorov-Smirnov test is based upon a simple idea. As the sample size n increases, the empirical distribution function $\hat{F}^{(n)}$ converges to the true underlying distribution function F. But this happens in such a way that the function $t \mapsto \sqrt{n}[\hat{F}^{(n)}(F^{-1}(t)) - t]$ converges in distribution to a Brownian bridge, tied down at 0 and 1. Hence,

$$\sqrt{n} \sup_{t \in [0,1]} [\hat{F}^{(n)}(F^{-1}(t)) - t] = \sqrt{n} \sup_{x \geq 0} [\hat{F}^{(n)}(x) - F(x)] \tag{5}$$

has approximately the distribution of the maximum of a Brownian bridge, which is a known distribution and can be looked up in tables. Precisely, if we hypothesise a distribution H, and observe that with a sample of size n,

$$\sqrt{n} \max_{x \geq 0} |\hat{F}^{(n)}(x) - H(x)| > 1.22 \tag{6}$$

then H is rejected at the 10% level of significance.

We are of course unable to apply directly the Kolmogorov-Smirnov test since we do not observe the empirical distribution function F_1 of X_1. However we can use the classification theorem described above to give a lower-bound estimate of the K-S distance $|\hat{F}^{(n)}(x) - H(x)|$. This is based on the following theorem,

Theorem 1 *Given an empirical sub-distribution function*

$$\hat{G}_1(x) = \frac{1}{n} \sum_{i=1}^{n} I(X_1^{(i)} \leq x, X_1^{(i)} < X_2^{(i)}) \tag{7}$$

and a distribution function H, let

$$D = \inf_{H_1 \in C^+(H)} \sup_x |\hat{G}_1(x) - H_1(x)|, \tag{8}$$

where $C^+(H)$ *is the class of functions* H_1 *such that* H_1 *and* $H - H_1$ *are nonnegative and nondecreasing (*$H = H_1 + (H - H_1)$ *is a nonnegative* comonotone representation *of* H*). Then for* any *distribution function* $\hat{F}$ *such that* $\hat{G}_1 \in C^+(\hat{F})$*,*

$$\sup_x |\hat{F}(x) - H(x)| \geq D \ . \tag{9}$$

In other words, whatever the (unseen) empirical distribution function $\hat{F}$ for X_1 is, its K-S distance to H is bounded below by D. Hence, if D has been calculated from the data and $\sqrt{n}D > 1.22$, then also $\sqrt{n}\sup_x |\hat{F}(x)-H(x)| \geq 1.22$ and H is rejected (conservatively) at the 10% significance level.

What makes this theorem very useful in practice is that, given an hypothesised distribution function H and an empirical sub-distribution function $\hat{G}_1$, the K-S distance D can be calculated by a fast and simple dynamic programming algorithm. The details can be found in Bedford and Meilijson [1].

4 An example of Quality Control

We now give an example to show how the test can work in practice in selected parametric cases. Non-parametric asymptotic analysis is still to be developed.

A machine with continuous lifetime X_1 is subject to quality control that detects failures that would have made the machine fail prior to age x_0. That is, if $X_1 \leq x_0$ then $X_2 = 0$ (the machine will not be used and its actual failure time will not be observed), while otherwise X_2 is anything exceeding X_1, such as $+\infty$ (the machine will fail at some age beyond x_0 and its lifetime will be observed).

Given a sample of n machines with m censored observations and $n - m$ observed failures, the empirical sub-distribution function $\hat{G}_2$ of X_2 is identically equal to $\frac{m}{n}$. The empirical lower Peterson bound for F_1, the (observed) empirical sub-distribution function $\hat{G}_1$ of X_1, is conveniently described in terms of its (unobserved) empirical distribution function $\hat{F}_1$ as $\hat{G}_1 = \max(0, \hat{F}_1 - \frac{m}{n})$, that is, it is equal to zero up to x_0 and is vertically parallel to $\hat{F}_1$ thereafter. The empirical upper Peterson bound for F_1, the empirical distribution function $\hat{\overline{F}}_1$ of $\min(X_1, X_2)$, is equal to $\frac{m}{n}$ up to x_0, and coincides with $\hat{F}_1$ thereafter.

The rules of the game are as follows: We have the data $\hat{G}_2, \hat{G}_1$ and $\hat{\overline{F}}_1$, but we don't know that the data was generated by the quality control procedure described above. (If we had known that, we would have known that the non-constant section of $\hat{\overline{F}}_1$ coincides on its domain of definition with the empirical distribution function $\hat{F}_1$.)

If we postulate that F_1 belongs to some family of distributions strictly ordered by stochastic inequality, it is clear that the upper Peterson bound is sharp. For the sake of concreteness, let us postulate that X_1 is exponentially distributed, i. e., $F_1(x) = 1 - e^{-\lambda x}$ for some $\lambda > 0$. Let us further assume that the data was generated with $\lambda = 1$.

The upper Peterson bound is thus sharp, so that the empirical upper Peterson bound implies asymptotically (as $n \to \infty$) that $\lambda \leq 1$. Letting $x_0 = 1$, the empirical lower Peterson bound yields asymptotically that $\lambda \geq 0.1355$, since these are the values of λ for which $1 - e^{-\lambda x} \geq e^{-1} - e^{-x}$ for all $x \geq 0$. However, the co-monotonicity restriction implies that the *density* $\lambda e^{-\lambda x}$ must be bigger or equal to the derivative e^{-x} of the lower Peterson bound, for all $x \geq x_0$. Elementary analysis shows that for $x_0 = 1$, the only feasible values of λ are $\lambda \geq 1$. Hence λ is *identifiable.*

The following theorem, generalising this example, shows that $x_0 = 1$ is the *borderline* of identifiability: if $x_0 \leq 1$ then $\lambda = 1$ is identifiable, while under a more efficient quality control cut-off ($x_0 > 1$), there is an interval $[\lambda_0, 1]$ of feasible values of λ, with $\lambda_0 = \lambda_0(x_0) < 1$.

Definition 2 *Let $\{F_\theta, \theta \in \Theta\}$ be a family of distributions ordered by monotone likelihood ratio, i.e., such that for every $\theta_0 < \theta_1 (\in \Theta)$, the function $\frac{dF_{\theta_1}(x)}{dF_{\theta_0}(x)}$ is strictly increasing. For an arbitrary $\theta \in \Theta$, the* calibrator *of the distribution F_θ, (or* inverse MLE*), is the point $x = x(\theta)$ such that if x was the single observation in a sample of size* 1 *then θ would have been the Maximum Likelihood estimate of this distribution's parameter. Equivalently, the calibrator of F_θ is the limit as $\theta' \to \theta$ of the point at which $\frac{dF_{\theta'}(x)}{dF_\theta(x)}$ crosses* 1.

For exponential-type families with density of the form $h(x)e^{\theta x - b(\theta)}$, the calibrator is equal to the mean. The calibrator is a "typical" point of a distribution in the context of a family of distributions containing the given one. We only present the following elegant property of this otherwise peculiar notion, without promoting its use. It states that distributions are identifiable if and only if their typical values escape censoring.

Theorem 2 *Let $\{F_\theta, \theta \in \Theta\}$ be a family of distributions ordered by monotone likelihood ratio. Suppose that for some $\theta_0 \in \Theta$ the random variable X_1 is F_{θ_0} - distributed and let (T, I) be generated by quality control censoring of X_1 with cut-off x_0. Then θ_0 is identifiable if and only if the cut-off x_0 is less than or equal to the calibrator of F_{θ_0}.*

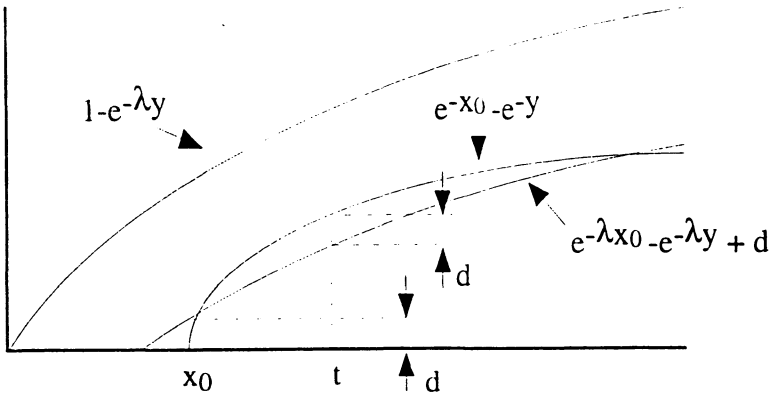

Figure 3: The best fit to the quality control subdistribution function

Proof:

As stated above, since monotone likelihood ratio order implies stochastic order, the upper Peterson bound sharply identifies θ as being less than or equal to the value θ_0. Let us consider now the co-monotone application of the lower Peterson bound. It states in this case that the density of F_θ must exceed the density of F_{θ_0} at all points to the right of x_0. If x_0 exceeds the calibrator of F_{θ_0}, there are values $\theta < \theta_0$ for which this is the case. If x_0 is below the calibrator, then there is an open interval around θ_0 where this *cannot* be the case except for θ_0 itself. In the borderline identifiable case where x_0 happens to coincide with the calibrator of F_{θ_0}, co-monotonicity allows for values of θ from θ_0 and up. □

In the light of the above proof, we see that under "solid" identifiability (i.e. when x_0 is strictly less than the calibrator) the co-monotone application of the lower Peterson bound locally identifies θ_0 by itself, while under "borderline" identifiability, it is only the joint action of both lower and upper bounds that identifies θ_0. Under these circumstances, it should come as no surprise that tests based on the lower bound may show pathologically slow convergence under borderline identifiability. We now return to the assumption of exponentiality to examine this question.

Consider the function $\hat{G}_1(y)$. It is equal to zero to the left of x_0 and is very close to $e^{-x_0} - e^{-y}$ to the right of x_0. Without loss of order of magnitude, we can assume that $\hat{G}_1$ is identically equal to this idealised function. Fix a value of λ below 1. The best vertical fit to $\hat{G}_1$ by a function growing at most as fast as $1 - e^{-\lambda y}$ is achieved by the function $\max(0, d + e^{-\lambda x_0} - e^{-\lambda y})$, with $d > 0$ chosen so that on the interval where this function is below $\hat{G}_1$, the maximal distance between the two is exactly equal to d, just as it is on the interval around x_0 where their order is reversed. See Figure 3. A short computation reveals that the point $t > 1$ where this maximal vertical distance is achieved is

$$t = \operatorname{argmin} \{(1 - e^{-y}) - (1 - e^{-\lambda y})\} = -\frac{\log \lambda}{1 - \lambda}. \tag{10}$$

Substituting this argument t as in

$$[(e^{-x_0} - e^{-y}) - (e^{-\lambda x_0} - e^{-\lambda y} + d)]|_{y=t} = d \tag{11}$$

we get, equivalently, letting

$$f(\lambda) = e^{-x_0} - e^{-\lambda x_0} - e^{\frac{\log \lambda}{1-\lambda}} + e^{\frac{\lambda}{1-\lambda} \log \lambda}, \tag{12}$$

that $f(\lambda)$ is equal to $2d$. Thus, a best fit (minimising d) can be obtained by finding a value of λ that minimises f. We now display formulas for the first three derivatives of f at $\lambda = 1$, as a function of x_0, and the value of each of these derivatives when the cut-off point x_0 is itself equal to 1.

$$\begin{aligned} f(1) &= 0 \\ f'(1) &= -(e^{-1} - x_0 e^{-x_0}) \end{aligned} \qquad \begin{aligned} f'(1)|_{x_0=1} &= 0 \\ f''(1)|_{x_0=1} &= 0 \\ f'''(1)|_{x_0=1} &= -\tfrac{3}{4e} \end{aligned} .$$

So, for $x_0 < 1$ we have that $(e^{-1} - x_0 e^{-x_0})(1 - \lambda) \approx 2d$, or

$$\lambda \approx 1 - \frac{2}{e^{-1} - x_0 e^{-x_0}} d \tag{13}$$

while for $x_0 = 1$ we have that $(-\frac{3}{4e})\frac{(1-\lambda)^3}{3!} \approx 2d$, or

$$\lambda \approx 1 - 3.517 d^{\frac{1}{3}}. \tag{14}$$

Since d is of order of magnitude $n^{-\frac{1}{2}}$, this rough computation yields that for x_0 close to 0 (no censoring) we should get that the left endpoint of the confidence interval for λ is approximately $1 - 5.44/\sqrt{n}$, for $x_0 = \frac{1}{2}$ we should get $1 - 31/\sqrt{n}$ and the coefficient of $1/\sqrt{n}$ should diverge as x_0 approaches 1, but for $x_0 = 1$ itself the order of magnitude is $1 - 3.517 n^{-\frac{1}{6}}$. In all cases, the right endpoint is the regular Kolmogorov - Smirnov right endpoint obtained via the Peterson upper bound, and its order of magnitude is $1 - 1/\sqrt{n}$.

These orders of magnitude ought to be validated by simulation, and we did. In fact, we performed the above analysis *because* of the extremely slow convergence we got for simulated data under x_0 equal or close to 1, ...after non-negligible efforts to find a bug in the program!

5 An example of independent censoring

Suppose that the lifetime T of a machine is exponentially distributed, there are two failure modes with failure times X_1, X_2, and its $\{\{1\},\{2\}\}$-valued cause of failure I is independent of T. Let the failure rate of T be θ and let $P(I = \{1\}) = p_1$. The unique independent model (X_1, X_2) for this observed data joint distribution makes X_1 and X_2 exponentially distributed with respective failure rates $p_1\theta$ and $(1-p_1)\theta$. However, within the rules of the game under which we know the joint distribution of the observed data (or have a random sample drawn from it) but not the dependence model that generated it, we have to identify the possible marginal distributions of X_1 and X_2. Suppose that we postulate X_1 to be marginally exponentially distributed. What is then the range of possible values of its failure rate λ? Since the sub-distribution function of X_1 is $G_1(x) = P(T \leq x)P(I = \{1\}) = (1 - e^{-\theta x})p_1$ and λ must satisfy $F_1'(x) = (1 - e^{-\lambda x})' \geq G_1'(x)$ for all $x > 0$, we have that $p_1\theta \leq \lambda \leq \theta$. The upper Peterson bound tells us that $\lambda \leq \theta$. However, the improvement of the upper Peterson bound by Bedford and Meilijson [1] alluded to in Section 2 rejects θ itself as a feasible value of λ, and we are left with the feasibility interval

$$p_1\theta \leq \lambda < \theta \,. \tag{15}$$

It is interesting to notice that the regular Peterson bound, without its co-monotone strenghtening, yields the interval $[p_1\theta, \theta]$ as well.

In this example, the independence model assesses λ as being equal to the lower bound $p_1\theta$ of all feasible values of λ. In other words, component 1 can be *at most as reliable* as this model claims. Optimism is a valuable property as long as it is recognised as such, but optimistic reliability assessments that are used as if they were typical values, can be dangerous. The following theorem strengthens the scope of this warning by showing that it holds in broader generality. Further evidence for the extent of this unsafe feature of the independence model is provided by Zheng and Klein [12]

> "Note that as the assumed strength of association increases the estimated survival function becomes smaller, so that an assumption of independence gives us overly optimistic estimates of the survival function."

Theorem 3 *Let $\{F_\theta, \theta \in \Theta\}$ be a family of distributions ordered by monotone likelihood ratio, to which the marginal distribution of X_1 is postulated to belong. Let X_1 be F_{θ_0} - distributed, with essential infimum x_0, and let (T, I) be generated as independent censoring of X_1 by a random variable X_2 with $P(X_2 > x_0) = 1$. Then θ_0 is the highest possible feasible value of θ.*

Proof:

For the derivative of the lower Peterson bound of F_θ we have

$$\underline{F}_1'(x) = F_{\theta_0}'(x)P(X_2 > x), \tag{16}$$

and by hypothesis $P(X_2 > x) \to 1$ as $x \searrow x_0$. Hence, feasible values of the marginal density $f_\theta(x)$ must exceed the corresponding value $f_{\theta_0}(x)$, for x sufficiently close to x_0. Since monotone likelihood ratio ordered families of distributions have their lower-end density values ordered in the opposite direction, the result follows. □

References

[1] T. Bedford and I. Meilijson. A characterisation of marginal distributions of (possibly dependent) lifetime variables which right censor each other. Preprint, Faculty of Technical Mathematics and Informatics, TU Delft, no. 93-116.

[2] R.M. Cooke. The total time on test statistic and age-dependent censoring. *Stat. and Prob. Let.*, **18**(5),1993.

[3] M. Crowder. On identifiability crisis in competing risk analysis. *Scand. J. Statist.***18**, 223-233, 1991.

[4] E.L. Kaplan and P. Meier. Non-parametric estimation from incomplete observations. *J. Amer. Statist. Assoc.* **53** 457-481, 1958.

[5] N. Langberg, F. Proschan and A.J. Quinzi. Converting dependent models into independent ones, preserving essential features. *Ann. Probability*, **6** 174-181, 1978.

[6] N. Langberg, F. Proschan and A.J. Quinzi. Estimating dependent life lengths, with applications to the theory of competing risks. *Ann. Statist.*, **9** 157-167, 1981.

[7] D.R. Miller. A note on independence of multivariate lifetimes in competing risk models. *Ann. Statist.*, **5**, 576-579, 1977.

[8] A. Nádas. On estimating the distribution of a random vector when only the smallest coordinate is observable. *Technometrics* **12**, 923-924, 1970.

[9] A. Nádas. The distribution of the identified minimum of a normal pair determines the distribution of the pair. *Technometrics* **13**, 201-202, 1971.

[10] A.V. Peterson. Bounds for a joint distribution function with fixed sub-distribution functions: Application to competing risks. *Proc. Nat. Acad. Sci. USA*, **73** (1), 11-13, 1976.

[11] A. Tsiatis. A non-identifiability aspect in the problem of competing risks. *Proc. Nat. Acad. Sci. USA*, **72** 20-22, 1975.

[12] M. Zhing and J.P. Klein. Estimates of marginal survival for dependent competing risks based on an assumed copula. Preprint, Department of Statistics, The Ohio State University, 1994.