COMPARING QUANTILE RESIDUAL LIFE FUNCTIONS BY CONFIDENCE BANDS

ALBA M. FRANCO-PEREIRA*, ROSA E. LILLO AND JUAN ROMO

Abstract

A quantile residual life function is the quantile of the remaining life of a surviving subject, as it varies with time. In this article we present a nonparametric method for constructing confidence bands for the difference of two quantile residual life functions. These bands provide evidence for two random variables ordering with respect to a quantile residual life order introduced in Franco-Pereira et al. (2010). A simulation study has been carried out in order to evaluate and illustrate the performance and the consistency of this new methodology. We also present applications to real data examples.

Keywords: quantile residual life, confidence bands.

* Universidad Carlos III de Madrid, Departament of Statistics, Facultad de Ciencias Sociales y Juridicas, C/ Madrid 126, 28903 Getafe (Madrid), e-mail: alba.franco@uc3m.es (Alba M. Franco-Pereira).

1   Introduction

The median residual life function (*merl*) represents the median additional time to failure given no failure by time t; more generally, the γ-quantile residual life function $q_{X,\gamma}(t)$ is defined as the γ-quantile of the remaining life given survival up to time t. That is, if lifelengths of the population are described by a random variable X with survival function $\overline{F}_X$, the median residual life function is defined by

$$merl(t) = \overline{F}_X^{-1}(0.5 \cdot \overline{F}_X(t)) - t, \quad t < u_X, \tag{1.1}$$

where F^{-1} denotes the left continuous version of the inverse of F:

$$F^{-1}(p) = \inf\{x : F(x) \geq p\}, \quad p \in (0,1),$$

and u_X is the right endpoint of its support. In general, if γ is a number between 0 and 1, the γ-quantile residual life function of X is defined by

$$q_{X,\gamma}(t) = F_X^{-1}(\gamma + \overline{\gamma}F_X(t)) - t, \quad t < u_X, \tag{1.2}$$

where $\overline{\gamma} = 1 - \gamma$. The quantile residual life function was first introduced by Haines and Singpurwalla (1974) as an alternative to the mean residual life function. The mean residual life function is a useful tool for analyzing important properties of X when it exists because it characterizes the distribution. However, it may not exist. Even when it exists it may have some practical shortcomings, especially in situations where the data are censored, or when the underlying distribution is skewed or heavy-tailed. In such cases, either the empirical mean residual life function cannot be calculated, or a single long-term survivor can have a marked effect upon its estimation, which will tend to be unstable due to its strong dependence on very long durations. The α-quantile residual life functions were studied in some detail by Arnold and Brockett (1983), Gupta and Langford (1984), Haines and Singpurwalla (1974), Joe and Proschan (1984a), Joe (1985) and Lillo (2005). Raja Rao et al. (2006) identified families of distributions for which simple expressions for the α-quantile residual life functions can be obtained.

In light of the extensive use of the quantile residual life functions in several areas of probability, statistics, and their applications, Franco-Pereira et al. (2010) introduce a new family of stochastic orders based on the pointwise comparison of these functions. In Jeong et al. (2008) is widely motivated the usefulness of comparing the quantile residual life functions; specially in the treatment of cancer when there exists a possibility of benefiting from a new secondary therapy to be given in the middle of the follow-up period and thus the Kaplan-Meier plots are not illustrative. They propose a score type test for the median residual life time to compare two groups at *any* fixed time point during the follow-up period. Their work improves the previous works of Wang and Hettmansperger (1990) who proposed a confidence interval approach to compare quantiles from failure time distributions under censoring. For the non-shift model, their methods involve the underlying density function. Su and Wei (1993) introduced a nonparametric test statistic to improve Wang and Hettmansperger's procedure. Both methods, however, were intended only for comparing the remaining lifetimes of patients at the origin of the follow-up period. Earlier, Berger et al. (1988) proposed a

modified test statistic based on Fligner and Rut's (1982) approach to compare two median residual lifetimes under censoring. However, unlike Su and Wei (1993), their approach also involves a nonparametric estimation of the probability density function of the failure time distribution under censoring, which is the major drawback in the median-based inference procedures for censored survival data. In Jung et at. (2009) the γ-quantile residual life function is associated at a specific time t_0 with selected covariates to develop a time-specific log-linear regression method on quantile residual lifetime. This work is useful to infer the remaining life years of a patient, given the genetic and/or environmental backgrounds.

In this paper we present an alternative method for comparing quantile residual life functions which is based on the construction of confidence bands for the difference of two quantile residual life functions. Although it is more common in the literature to compare quantile residual life functions with $\gamma = 0.5$, if the objective of a study is inferring about the population with longer survival, high quantiles such as 80 or 90 percentile, rather than the median, would be more appropriate. The comparison of quantile residual life functions is also useful in Finance since a common measurement of the risk is the *Value at Risk*, which is defined as the γ-quantile of the loss distribution for some prescribed confidence level $\gamma \in (0, 1)$. For more applications of these stochastic orderings see Franco-Pereira et al. (2010).

Joe and Proschan (1984b) and Cheng (1985) proposed tests for comparing hazard rate functions and quantile residual life functions, respectively. These tests, as well as many other tests proposed to compare functions, checked the null hypothesis that two functions were equal for all t, versus the alternative that one dominated the other for all t. These models did not account for the realistic possibility that the functions cross. A test designed only to test the null hypothesis of equality may have a large probability of rejecting this null hypothesis for two populations whose functions cross. Rejection of the null hypothesis by such a test may be interpreted as evidence that one function dominates the other only if the possibility of crossing functions can be eliminated a priori. In that sense, the tests proposed by Berger et al. (1988) to compare mean residual life functions and median residual life functions included the possibility that the functions cross. The null hypothesis in their tests is that the two functions are ordered in an interval (which may be the whole support) versus the alternative that they are not.

The bands we construct evidence whether two random variables are close with respect to the quantile residual life order or not in the following sense. If the random variables are ordered, we will expect that the whole band lies below or above the x-axis. Otherwise, we will expect that at least one of the limits of the band intersects the x-axis or each limit of the band is at a different side of the x-axis. One of the advantages of the proposed procedure is that, in case that the variables are not ordered with respect a the quantile residual life order, it allow us to detect the moment at which the functions cross. That is, the shape of the band evidences the crossing effect when it exists. The method is also valid when there are censored data. Moreover, we do not need all the data to apply our methodology.

Given two random samples, we estimate their quantile residual life functions and their bootstrapped counterparts, compute the difference between them and apply statistical depth for functions as a criteria for constructing the bands. Statistical depth is a natural tool to analyze functional data aspects. It has been introduced to measure the 'centrality' or the

'outlyingness' of an observation with respect to a given dataset or a population distribution (see Ramsay and Silverman (2005)). The notion of depth was first considered for multivariate data to generalize order statistics, ranks, and medians to higher dimensions. However, direct generalization of current multivariate depths to functional data often leads to either depths that are computationally intractable or depths that do not take into account some natural properties of the functions, such as shape. For that reason, some specific definitions of depth for curves have been introduced, to provide us with a criteria to order the sample curves from the center-outward (from the deepest to the most extreme). See, for example, Vardi and Zhang (2000), Fraiman and Muniz (2001), López-Pintado and Romo (2005), Cuevas, Febrero and Fraiman (2007), Cuevas and Fraiman (2009) and López-Pintado and Romo (2009). We will use this last depth concept to choose the $(1 - \alpha)$ more central curves, which will be illustrated in Section 2.

This paper is organized as follows. Our methodology to construct the confidence bands is explained in Section 2. In Section 3 we describe the simulation study we have carried out and in Section 4 we apply the new method to real data examples.

2   Our methodology

In this section we describe the methodology we have considered to define a nonparametric method for constructing confidence bands for the difference of two quantile residual life functions. We apply bootstrap techniques and statistical depth as a criteria for constructing the bands. In the sequel we detail the methodology.

For a sample size n, let $X_1, X_2, \ldots, X_n$ be independent random variables with common distribution function F_X, and let $X_{1:n}, X_{2:n}, \ldots, X_{n:n}$ be the corresponding order statistics. Besides the sample distribution function $F_n(t) = n^{-1}\sharp\{k : X_k \leq t, 1 \leq k \leq n\}$, we consider also the sample quantile function

$$Q_n(y) = \begin{cases} X_{k:n} & \frac{k-1}{n+1} < y \leq \frac{k}{n+1}, \quad k = 1, \ldots, n, \\ X_{n:n} & \frac{n}{n+1} < y \leq 1. \end{cases}$$

The natural empirical counterpart of $q_{X,\alpha}$, the sample α-quantile residual life at $t > 0$, is

$$q_{X,n,\alpha}(t) = Q_n(\bar{\alpha}\bar{F}_{X,n}(t)) - t, \quad t < u_X, \tag{2.1}$$

where $\bar{F}_{X,n} = 1 - F_{X,n}$ denotes the empirical survival function of X.

Note that $q_{X,n,\alpha}(t)$ is a piecewise linear function with jump discontinuities. Assuming that F_X has a continuous positive density function f_X such that $\inf_{\{0 \leq y \leq 1\}} f_X(F_X^{-1}(y)) > 0$, then the sample α-quantile residual life function is a strongly uniformly consistent estimator. This result can be found in Corollary 1.4.1. in Csörgö (1983) and the following paragraph.

Now, let Y be another random variable and let $Y_1, Y_2, \ldots, Y_m$ be an i.i.d. random sample from Y. Given $\gamma \in (0, 1)$, plots of the empirical γ-quantile residual life functions of X and

Y can give an indication of the plausibility of whether the γ-quantile residual life functions of X and Y are ordered; that is, if X and Y are ordered according to the γ-quantile residual life order.

Since $q_{X,n,\gamma}$ and $q_{Y,m,\gamma}$ are strong consistent estimators of $q_{X,\gamma}$ and $q_{Y,\gamma}$, respectively,

$$q_{Y,m,\gamma}(t) - q_{X,n,\gamma}(t) \to q_{Y,\gamma}(t) - q_{X,\gamma}(t),$$

almost surely, as $n, m \to \infty$.

For constructing the band we follow a three-step algorithm. Let B be the bootstrap size, $\alpha \in (0, 1)$ be the confidence level and $\gamma \in (0, 1)$ be the fixed quantile.

(i) Consider the bootstrap replications from $X_1, \ldots, X_n$ and $Y_1, \ldots, Y_m$. We denote them by $X_1^{*b}, \ldots, X_n^{*b}$ and $Y_1^{*b}, \ldots, Y_m^{*b}$, respectively, for $b = 1, \ldots, B$.

(ii) For every $b = 1, \ldots, B$, compute the empirical γ-quantile residual life functions which is associated to $X_1^{*b}, \ldots, X_n^{*b}$ and $Y_1^{*b}, \ldots, Y_m^{*b}$. We denote them by $q_{X,n,\gamma}^{*b}$ and $q_{Y,m,\gamma}^{*b}$, respectively. Then consider
$$q_b^* = q_{Y,m,\gamma}^{*b} - q_{X,n,\gamma}^{*b}.$$

(iii) Order the sample curves q_b^*, $b = 1, \ldots, B$, from inner to outer using any concept of depth for functions and take the band given by the $(1 - \alpha) \cdot 100\%$ deepest curves.

The convex hull for these $(1 - \alpha) \cdot 100\%$ deepest curves constitutes the confidence band.

Remark 2.1. If there exist censored data in the sample, we can consider the quantile residual life estimator under censoring proposed in Csörgő (1987) and follow the bootstrap mechanism approach proposed by Efron (1981). The estimator of Csörgő (1987) is based on the following estimation of the survival function of X

$$\bar{F}_{X,n}(x) = \frac{N^+(x) + 1}{n + 1} \prod_{j=1}^{n} \left(\frac{2 + N^+(Z_j)}{1 + N^+(Z_j)} \right)^{I_{\{\delta_j = 0, Z_j \leq x\}}}, \tag{2.2}$$

where $N^+(x)$ denotes the number of censored and uncensored observations greater than x. Right-censored data is of the form $\{(x_1, d_1), \ldots, (x_n, d_n)\}$, where x_j is the jth observation, censored or not, and

$$d_j = \begin{cases} 1 & \text{if } j \text{ is uncensored,} \\ 0 & \text{if } j \text{ is censored.} \end{cases}$$

We draw a bootstrap sample $(X_1^*, D_1^*), (X_2^*, D_2^*), \ldots, (X_n^*, D_n^*)$ by independent sampling n times with replacement from $(X_1, D_1), (X_2, D_2), \ldots, (X_n, D_n)$, putting mass $1/n$ at each point. This form of the bootstrap requires only that the observed pairs (x_i, d_i) are i.i.d. observations from a distribution F_X on $\mathbb{R} \times \{0, 1\}$.

The notion of depth was first considered for multivariate data to generalize order statistics, ranks, and medians to higher dimensions. Several depth definitions for multivariate data have been proposed and analyzed by Mahalanobis (1936), Tukey (1975), Oja (1983),

4

Liu (1990), Singh (1991), Fraiman and Meloche (1999), Vardi and Zhang (2000), Koshevoy and Mosler (1997) and Zuo (2003). Direct generalization of current multivariate depths to functional data often leads to either depths that are computationally intractable or depths that do not take into account some natural properties of the functions, such as shape. For that reason several specific definitions of depth for functional data were introduced. See, for example, Vardi and Zhang (2000), Fraiman and Muniz (2001), López-Pintado and Romo (2005), Cuevas et al. (2007), Cuevas and Fraiman (2009) and López-Pintado and Romo (2009). The definition of depth for curves provides us with a criteria to order the sample curves from the center-outward (from the deepest to the most extreme).

In the third step of the algorithm, we will use one of the statistical depths introduced in López-Pintado and Romo (2009), the modified band depth. One of the advantages of this depth is that it follows a graph-based approach.

Let $x_1(t), \ldots, x_n(t)$ be a collection of real functions. Although the following ideas can be introduced for more general functional observations, we will restrict the exposition to functions in the space $C(I)$ of real continuous functions on the compact interval I. The graph of a function x is the subset of the plane $G(x) = \{(t, x(t)) : t \in I\}$ and the band in $\mathbb{R}^2$ delimited by the curves $x_{i_1}, \ldots x_{i_k}$ is

$$B(x_{i_1}, \ldots x_{i_k}) = \{(t, y) : t \in I, \min_{r=1,\ldots,k} x_{i_r}(t) \leq y \leq \max_{r=1,\ldots,k} x_{i_r}(t)\} =$$

$$= \{(t, y) : t \in I, y = \beta \min_{r=1,\ldots,k} x_{i_r}(t) + (1 - \beta) \max_{r=1,\ldots,k} x_{i_r}(t), \beta \in [0,1]\}.$$

Figure 1 (a) presents the band $B(x_1, x_2)$ given by two curves whereas Figure 1 (b) presents the band $B(x_1, x_2, x_3)$ given by three curves.
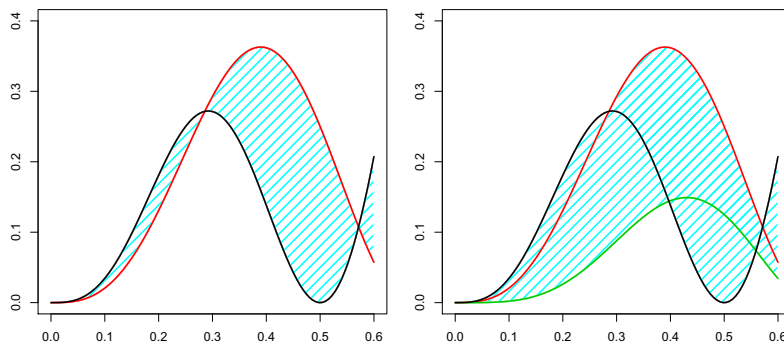

Figure 1: (a) Band defined by two curves x_1 (black) and x_2 (red). (b) Band given by three curves x_1 (black), x_2 (red) and x_3 (green).

For any function x in $x_1, \ldots, x_n$, and a fixed j value with $2 \leq j \leq n$, the quantity

$$BD_n^{(j)}(x) = \binom{n}{j}^{-1} \sum_{1 \leq i_1 < i_2 < \ldots < i_j \leq n} I_{\{G(x) \subseteq B(x_{i_1}, x_{i_2}, \ldots x_{i_j})\}},$$

expresses the proportion of bands $B(x_{i_1}, x_{i_2}, ...x_{i_j})$ determined by j different curves $x_{i_1}, x_{i_2}, ...x_{i_j}$ containing the whole graph of x. ($I_{\{A\}}$ is one if A is true, and zero otherwise).

Now, let J be a fixed value with $2 \le J \le n$. For functions $x_1, ...x_n$, the *band depth* of any of these curves x is

$$BD_{n,J}(x) = \sum_{j=2}^{J} BD_n^{(j)}(x).$$

If $X_1, \ldots, X_J$ are independent copies of the stochastic process X generating the observations $x_1, \ldots, x_n$, the corresponding population versions are $BD^{(j)}(x, P) = P\{G(x) \subset B(X_1, \ldots, X_j)\}$ and

$$BD_J(x, P) = \sum_{j=2}^{J} BD^{(j)}(x, P) = \sum_{j=2}^{J} P(G(x) \subset B(X_1, \ldots, X_j)).$$

López-Pintado and Romo (2009) recommend considering the definition of band depth with $J = 3$ for several reasons: (1) when J is larger than 3 the index $BD_{n,J}$ can be computationally intensive, (2) bands corresponding to large values of J do not resemble the shape of any of the curves from the sample, (3) the band depth induced order is very stable in J, and (4) the band depth with $J = 2$ is the easiest to compute but, if two curves cross over, the band delimited by them is degenerated in a point and, with probability one, no other curve will be inside this band.

Instead of considering the indicator function, a more flexible definition was introduced by measuring the set where the function is inside the band. For any of the functions x in $x_1, \ldots, x_n$ and for $2 \le j \le n$, let

$$A_j(x) \equiv A(x; x_{i_1}, \ldots, x_{i_j}) \equiv \{t \in I : \min_{r=i_1,...,i_j} x_r(t) \le x(t) \le \max_{r=i_1,...,i_j} x_r(t)\}$$

be the set where the function x is in the band determined by the observations $x_{i_1}, x_{i_2}, \ldots, x_{i_j}$. If λ is the Lebesgue measure on I, $\lambda_r(A_j(x)) = \lambda(A_j(x))/\lambda(I)$ gives the 'proportion of time' that x is in the band. Now, for $2 \le j \le n$,

$$MBD_n^{(j)}(x) = \binom{n}{j}^{-1} \sum_{1 \le i_1 < i_2 < ... < i_j \le n} \lambda_r(A(x; x_{i_1}, \ldots, x_{i_j}))$$

is a more flexible version of $BD_n^j(x)$: if x is always inside the band, the value $\lambda_r(A_j(x))$ is one as in the previous notion of depth.

Figure 2 illustrates the 'proportion of time' that the curve x_1 (black) is inside the bands determined by different curves in the sample.

Now, let J be a fixed value with $2 \le J \le n$. For functions $x_1, \ldots, x_n$, the *modified band depth* of any of these curves x is
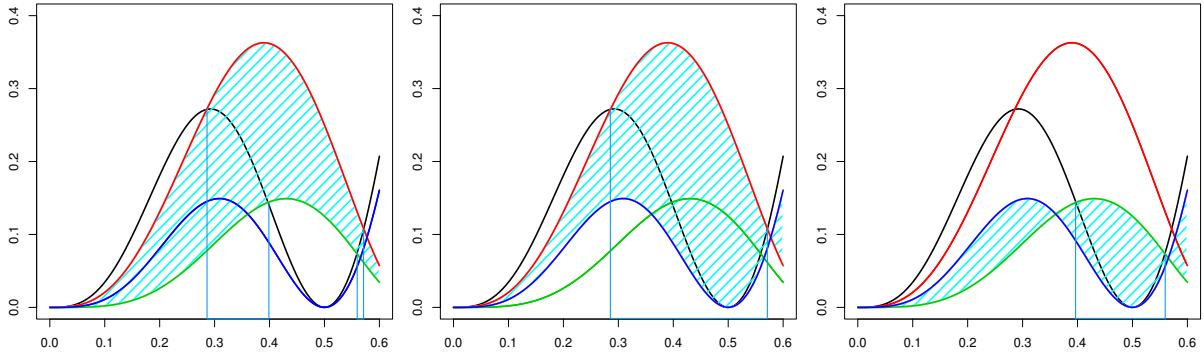
$$MDB_{n,J}(x) = \sum_{j=2}^{J} MDB_n^{J}(x).$$

Figure 2: Illustration of the 'proportion of time' that x_1 (black) is in the band determined by (a) x_2 (red) and x_3 (green) (b) x_2 (red) and x_4 (blue) (c) x_3 (green) and x_4 (blue).

The population version of the modified band depth is $MDB_J(x) = \sum_{j=2}^{J} MDB^{(j)}(x)$, where $MDB^{(j)}(x) = E[\lambda_r(A(x; X_1, X_2, \ldots, X_j))]$.

It is straightforward to check that in the univariate case, the band depth and the modified band depth coincide. Moreover, the ordering induced in that case does not depend on J. López-Pintado and Romo (2009) recommend considering the definition of modified band depth with $J = 2$ because it is computationally fast and the order induced is very stable in J, and (contrary to the band depth) it provides reasonable orders even if many curves from the sample cross over.

The band depth is more dependent on the curves shape and more restrictive than the modified version, providing frequent ties (several curves with the same depth). The modified band depth relies more on the magnitude or size of the curves that on their shape. Another relevant difference between them is their behavior for curves leaving the center of the sample only for a short interval (i.e., remaining in the interior of the sample almost all the time, but taking extreme values in short subintervals): the modified band depth can still be large for them but the band depth will always take small values on these curves.

Using the band depth or the modified band depth depends on the kind of functions being analyzed and the objectives to be checked. If the curves are very irregular, it is convenient to use the modified band depth because it avoids having too many depth ties and there might not be a representative 'shape'; however, the band depth is more adequate if the curves are smooth and the goal is the most representative curve in terms of shape (not magnitude).

3 Simulation study

A simulation study has been carried out in order to evaluate the performance and to illustrate the consistency of our methodology. As we will see, the bootstrap procedures play a central role in the methods studied here. The computer codes written in R can be found at **http://halweb.uc3m.es/esp/Personal/personas/afranco/eng/Software.html**.

As we pointed out in Section 1, if two random variables are ordered with respect to a γ-quantile residual life order, then the difference of the γ-quantile residual life functions associated to the variables must be either positive or negative in the whole support. Based on this idea, we have developed a nonparametric method for constructing confidence bands for the difference of two quantile residual life functions as it was explained in Section 2. These bands will provide us with an evidence of whether two random variables are close with respect to a quantile residual life order or not, in the following sense. If the random variables are ordered, we will expect that the whole band lies below or above the x-axis. Otherwise, we will expect that at least one of the limits of the band intersects the x-axis or each limit of the band is at a different side of the x-axis.

We have examined several examples through simulation. The random variables we have considered first follow Pareto distributions. The reason is that, for any $\gamma \in (0,1)$, the γ-quantile residual life function of a Pareto distribution is a line for $t \geq 0$ and for $t < 0$. Therefore, if we compare the γ-quantile residual life functions of two Pareto distributions there are only two possible situations: either their γ-quantile residual life functions are parallel for all $t \geq 0$ or they intersect. In the first case we conclude that X and Y are ordered in the γ-quantile residual life order and in the second case they are not.

Let X have the Pareto distribution

$$F_X(t) = 1 - \left(\frac{\rho_X}{\rho_X + t}\right)^{\nu_X}, \quad t \geq 0,$$

where $\rho_X > 0$ and $\nu_X > 0$. We denote it $X \sim P(\rho_X, \nu_X)$. Then, for any $\gamma \in (0,1)$,

$$q_{X,\gamma}(t) = \begin{cases} ((1-\gamma)^{-1/\nu_X} - 1)\rho_X - t, & t < 0; \\ ((1-\gamma)^{-1/\nu_X} - 1)(\rho_X + t), & t \geq 0. \end{cases}$$

See Franco-Pereira et al. (2010). Now, if $Y \sim P(\rho_Y, \nu_Y)$, with $\rho_Y > 0$ and $\nu_Y > 0$, it follows that

$$X \leq_{\alpha\text{-rl}} Y \iff \begin{cases} \nu_Y \leq \nu_X \quad \text{and} \\ \frac{(1-\alpha)^{-1/\nu_X}-1}{(1-\alpha)^{-1/\nu_Y}-1} \leq \frac{\rho_Y}{\rho_X}. \end{cases}$$

The following equivalence holds for the Pareto case:

$$X \leq_{\gamma-rl} Y \Leftrightarrow X \leq_{\beta-rl} Y, \text{ for all } \beta \in (0,1) \Leftrightarrow X \leq_{hr} Y,$$

where $\leq_{hr}$ denotes the hazard rate order. The second inequality always holds (see Franco-Pereira et al. (2010)).

From Example 3.1 to Example 3.6 we have considered six pairs of Pareto random variables where X is always the same and the Y_i vary for $i = 1, \ldots, 6$, and are ordered with respect to any quantile residual life order. In particular, they are ordered with respect to the median residual life order. The median residual life functions of these variables are represented in Figure 3. It can be seen that the distance between the median residual life function of X and the median residual life function of Y_i, $i = 1, \ldots, 6$ increases as i increases.

The pairs of variables from Example 3.7 to Example 3.10 are not ordered with respect to any γ-quantile residual life order. In particular, they are not ordered for $\gamma = 0.5$. The median residual life functions of these variables are represented in Figure 4.

8

Example 3.1.
$$\begin{cases} X & \sim P(10, 10) \\ Y_1 & \sim P(20, 10) \end{cases}$$

Example 3.2.
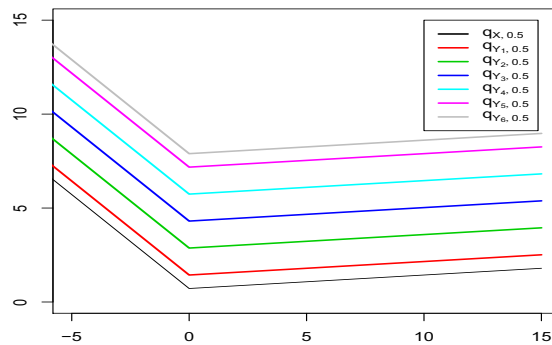$$\begin{cases} X & \sim P(10, 10) \\ Y_2 & \sim P(40, 10) \end{cases}$$

Example 3.3.
$$\begin{cases} X & \sim P(10, 10) \\ Y_3 & \sim P(60, 10) \end{cases}$$

Example 3.4.
$$\begin{cases} X & \sim P(10, 10) \\ Y_4 & \sim P(80, 10) \end{cases}$$

Example 3.5.
$$\begin{cases} X & \sim P(10, 10) \\ Y_5 & \sim P(100, 10) \end{cases}$$

Example 3.6.
$$\begin{cases} X & \sim P(10, 10) \\ Y_6 & \sim P(110, 10) \end{cases}$$

Figure 3: Median residual life functions of the variables from Example 3.1 to Example 3.6.

Example 3.7.
$$\begin{cases} X_7 & \sim P(10, 10) \\ Y_7 & \sim P(1, 5) \end{cases}$$
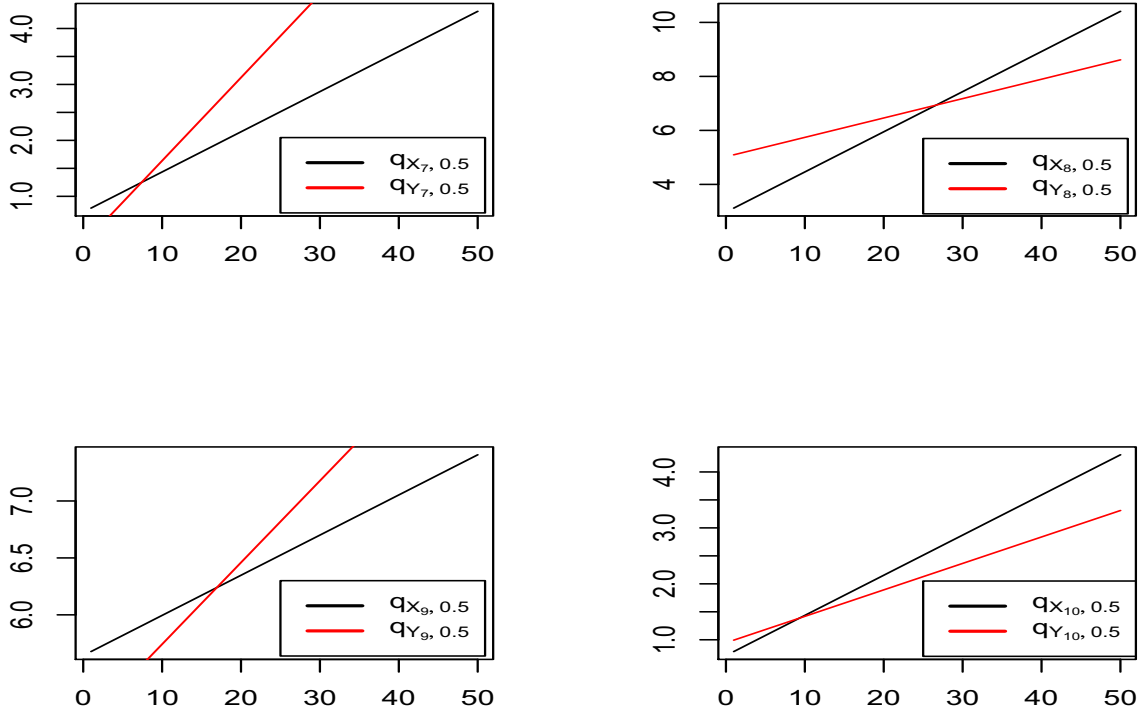
Figure 4: Median residual life functions of the variables from Example 3.7 to Example 3.10.

Example 3.8.

$$\begin{cases} X_8 & \sim P(20, 5) \\ Y_8 & \sim P(70, 10) \end{cases}$$

Example 3.9.

$$\begin{cases} X_9 & \sim P(160, 20) \\ Y_9 & \sim P(70, 10) \end{cases}$$

Example 3.10.

$$\begin{cases} X_{10} & \sim P(10, 10) \\ Y_{10} & \sim P(20, 15) \end{cases}$$

We have fixed $B = 1000$ the bootstrap size, $\gamma = 0.5$ the quantile we want to compare and $\alpha = 0.1$ the confidence level. Although all the procedure is valid for any definition of depth for curves, we have considered the modified band depth that we have explained in Section 2 with $J = 2$.

The 90%-confidence bands for the difference of the median residual life functions of the variables in the first six examples are shown in Figure 5 to Figure 10. In the left graph of each figure, the $B = 1000$ bootstrap estimations of the difference of the median residual life functions together with the confidence band is shown. In the right graph, only the band is

shown. The 90%-confidence bands for the difference of the median residual life functions of the variables in the last four Pareto examples are shown in Figure 11 to Figure 14.
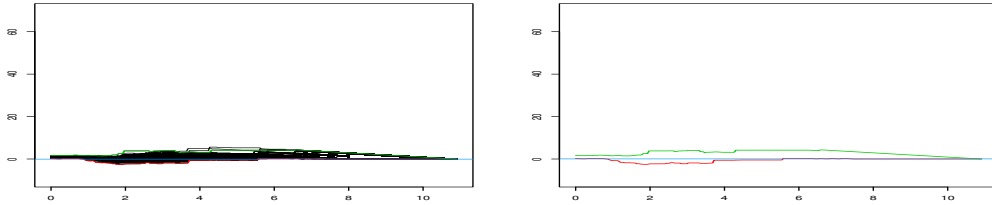
Figure 5: 90%-confidence band for $q_{Y_1,0.5} - q_{X,0.5}$.
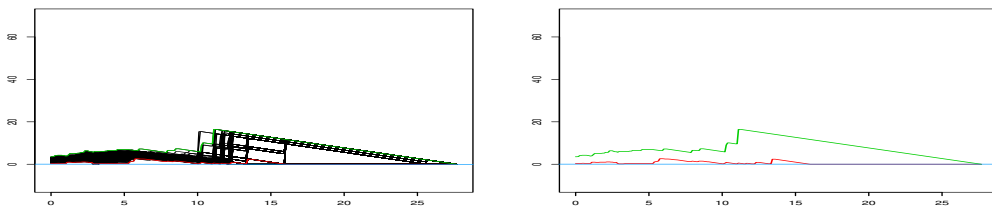
Figure 6: 90%-confidence band for $q_{Y_2,0.5} - q_{X,0.5}$.
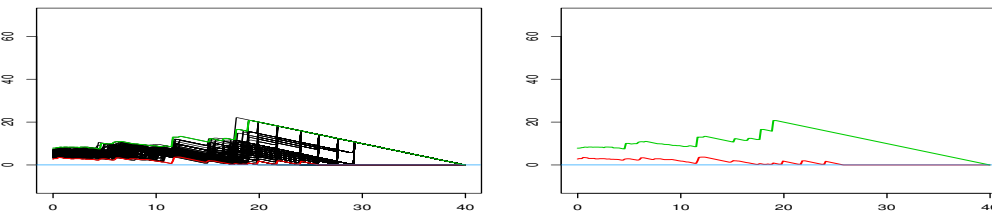
Figure 7: 90%-confidence band for $q_{Y_3,0.5} - q_{X,0.5}$.

In Figures 5 to 10, which correspond to the Pareto examples where the variables are ordered in the sense of the median residual life order (Examples 3.1 to 3.6), except when the difference of the variables is very small (Example 3.1), all the region delimited by the band lies above the x-axis. In Figure 5, in which the variables involved are ordered but close with respect to the median residual life order (Example 3.1), it is seen that the lower limit of the band crosses the x-axis or the x-axis is contained in the band. However, as the difference between the two median residual life functions increases, the distance of the lower limit of the band to the x-axis also increases. This fact shows the coherence of these bands.

In Figures 11 to 14, where the Pareto variables are not ordered in the sense of the median residual life order (Examples 3.7 to 3.10), the x-axis either cross any of the limits of the band or is contained inside the band. Actually, the proposed procedure in this case, where the variables are not ordered with respect to the quantile residual life order, allows us to detect the moment at which the functions cross.
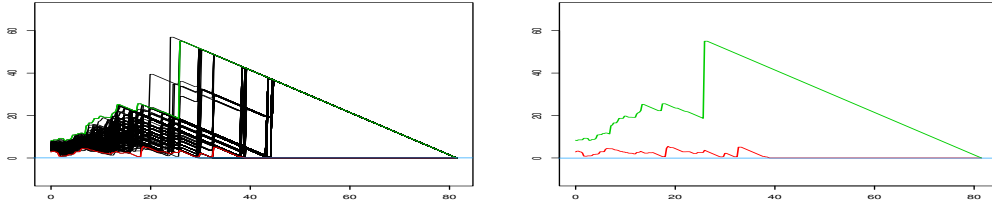
11

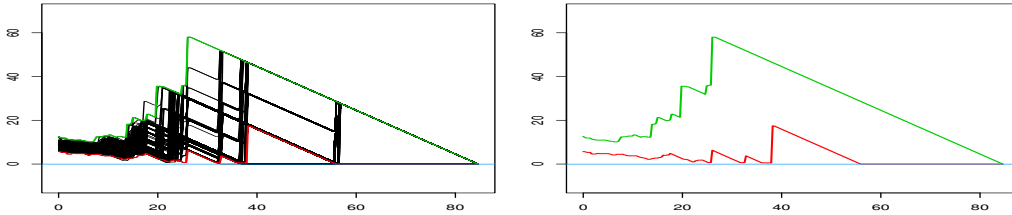Figure 8: 90%-confidence band for $q_{Y_4,0.5} - q_{X,0.5}$.

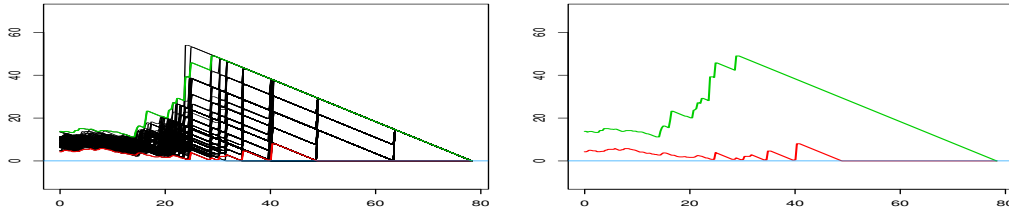Figure 9: 90%-confidence band for $q_{Y_5,0.5} - q_{X,0.5}$.

Figure 10: 90%-confidence band for $q_{Y_6,0.5} - q_{X,0.5}$.

We have also considered six examples with variables following Weibull distributions, which are very often used in reliability. Ecologists have long been interested in modelling survivorship of biological units such as plants and animals. The Weibull distribution has been noted to be one of the most suitable distributions in this context (Pinder et al. (1978)). There are number of reasons for its wide applicability. Firstly, the Weibull distribution can accommodate the three commonly encountered hazard or age-specific instantaneous mortality curves: constant, increasing, and decreasing. Secondly, the statistical methodology of the Weibull distribution is quite simple.

The quantile residual life function of the Weibull distribution is decreasing, therefore, it is very useful to model aging notion. See Lai and Xie (2006) for more details.

Let X has the Weibull distribution, then its distribution function is given by

$$F_X(t) = 1 - \exp\{-(\lambda_X t)^{m_X}\}, \quad t \geq 0,$$
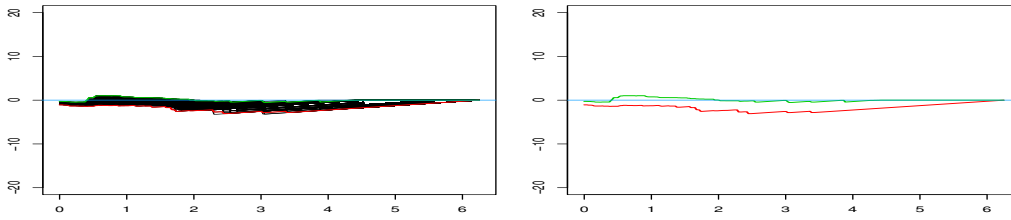
12

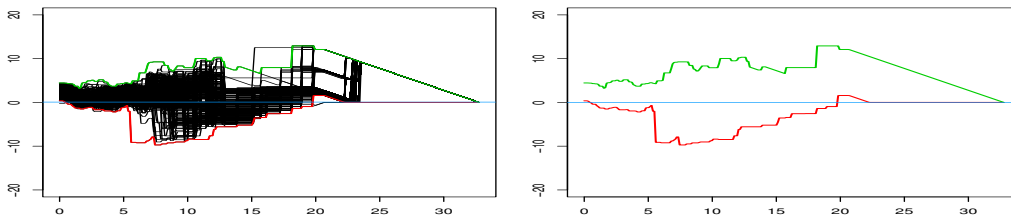Figure 11: 90%-confidence band for $q_{Y_7,0.5} - q_{X,0.5}$.

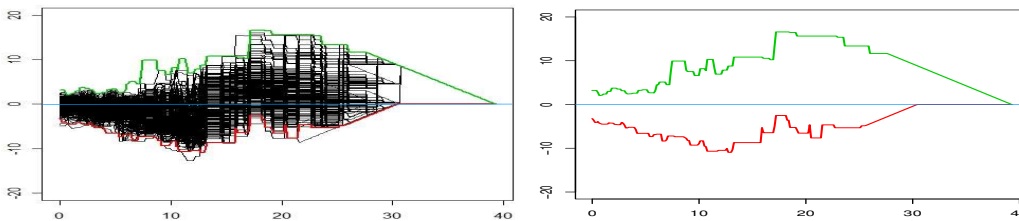Figure 12: 90%-confidence band for $q_{Y_8,0.5} - q_{X,0.5}$.

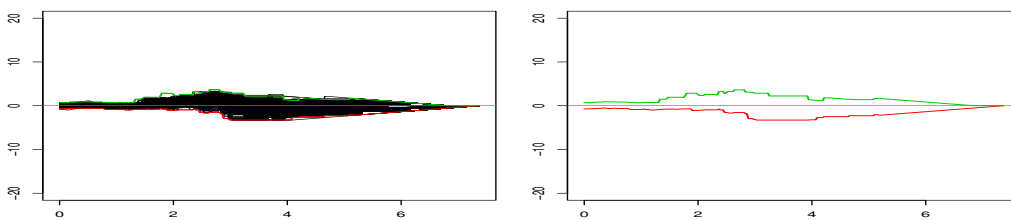Figure 13: 90%-confidence band for $q_{Y_9,0.5} - q_{X,0.5}$.

Figure 14: 90%-confidence band for $q_{Y_{10},0.5} - q_{X,0.5}$.

where $m_X > 0$ and $\lambda_X > 0$. We denote it $X \sim W(m_X, \lambda_X)$. Then, for any $\gamma \in (0,1)$,

$$q_{X,\gamma}(t) = \frac{[(\lambda_X t)^{m_X} - \log(1-\gamma)]^{\frac{1}{m_X}}}{\lambda_X} - t, \quad \text{for all } t.$$

It can be seen that if Y is another random variable such that $Y \sim W(m_Y, \lambda_Y)$ with

13

$m_X = m_Y$ and $\lambda_X \geq \lambda_Y$, then their γ-quantile residual life functions are ordered for all $\gamma \in (0,1)$. That is, X and Y are ordered in any quantile residual life order.

From Example 3.11 to Example 3.16 we have considered six pairs of Weibull random variables where X^* is always the same and the Y_i^* vary for $i = 1, \ldots, 6$, and that are ordered with respect to any quantile residual life order. In particular, they are ordered with respect to the median residual life order. The median residual life functions of these variables are represented in Figure 15. It can be seen that the distance between the median residual life function of X^* and the median residual life function of Y_i^*, $i = 1, \ldots, 6$ increases as i increases (as λ_{X^*} gets closer to $\lambda_{Y_i^*}$).

Example 3.11.
$$\begin{cases} X^* & \sim W(2,7) \\ Y_1^* & \sim W(2,6) \end{cases}$$

Example 3.12.
$$\begin{cases} X^* & \sim W(2,7) \\ Y_2^* & \sim W(2,5) \end{cases}$$

Example 3.13.
$$\begin{cases} X^* & \sim W(2,7) \\ Y_3^* & \sim W(2,4) \end{cases}$$

Example 3.14.
$$\begin{cases} X^* & \sim W(2,7) \\ Y_4^* & \sim W(2,3) \end{cases}$$

Example 3.15.
$$\begin{cases} X^* & \sim W(2,7) \\ Y_5^* & \sim W(2,2) \end{cases}$$

Example 3.16.
$$\begin{cases} X^* & \sim W(2,7) \\ Y_6^* & \sim W(2,1) \end{cases}$$

In Figure 16 to Figure 21, the 90%-confidence bands for the difference of the median residual life functions of the variables in the Weibull examples are shown. Again, except when the difference of the variables is very small (Example 3.11), all the region that is delimited by the limits of the band lies above the x-axis.

In Figure 16, where the variables involved are ordered but close with respect to the median residual life order (Example 3.11), it is seen that the lower limit of the band crosses the x-axis or the x-axis is contained in the band. However as the difference between the two median residual life functions increases, the distance of the lower limit of the band to the x-axis also increases. This fact, again, shows the coherence of these bands.
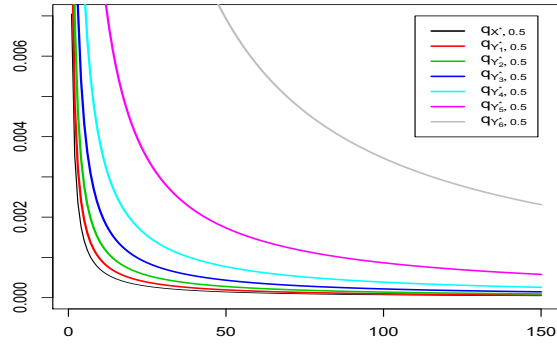
14

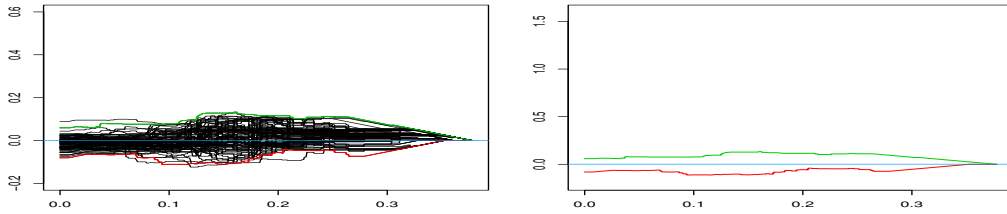Figure 15: Median residual life functions of the variables from Example 3.11 to Example 3.16.

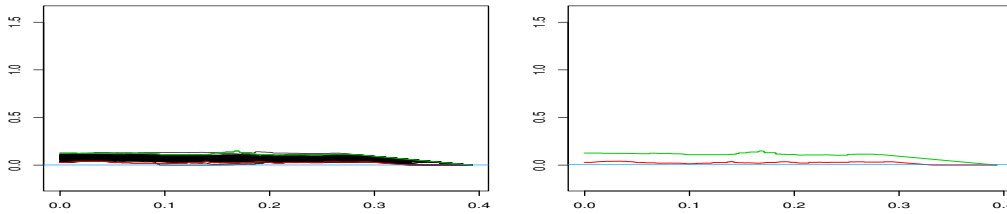Figure 16: 90%-confidence band for $q_{Y_1^*,0.5} - q_{X^*,0.5}$.

Figure 17: 90%-confidence band for $q_{Y_2^*,0.5} - q_{X^*,0.5}$.
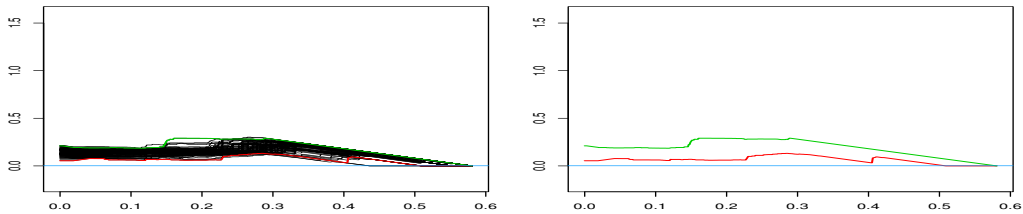

Figure 18: 90%-confidence band for $q_{Y_3^*,0.5} - q_{X^*,0.5}$.

Remark 3.17. We have computed the confidence band for the difference of two median residual life functions. However, as we have already explained, the procedure is valid for
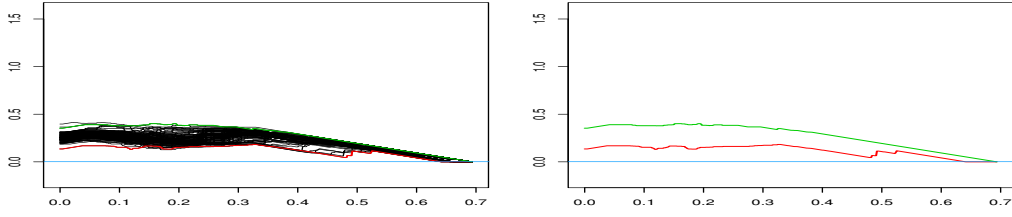
15

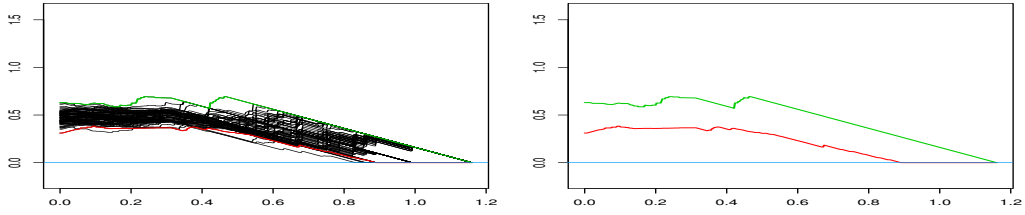Figure 19: 90%-confidence band for $q_{Y_4^*,0.5} - q_{X^*,0.5}$.

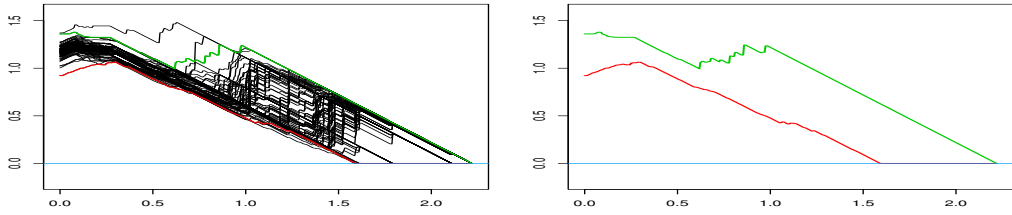Figure 20: 90%-confidence band for $q_{Y_5^*,0.5} - q_{X^*,0.5}$.

Figure 21: 90%-confidence band for $q_{Y_6^*,0.5} - q_{X^*,0.5}$.

any quantile $\gamma \in (0,1)$, and the graphs of the band continue being consistent with the order between the functions.

As we have illustrated with this examples, the confidence bands for the difference of two quantile residual lifetimes provide us with a criteria of whether two random variables are close or not with respect to a quantile residual life order, or allow us to compare quantile residual life functions in a given interval.

4   Application to real data examples

Failure time analysis (FTA) addresses data of the form 'time until an event occurs'. The survival times of medical patients or industrial products have been the usual subjects of FTA, but also data from a wide variety of ecological studies may be cast in these terms, including survival times of organisms or part of organisms and times until certain behaviours are exhibited.

In the biomedical context, FTA has been also called 'survival analysis' since the event is commonly the death of a patient, so the time until the death is the survival time. Failure time analysis accommodates 'censored data'. Censored data points are those in which the event was not observed, perhaps because the study ended before the event happened to some of the individuals under observation or because some of the individuals were lost track of before the event occurred during the study. For these censored data points, the actual time of occurrence is not known. Instead we know a minimum length of time during which the event did not occur. Failure time analysis allows use of such censored data for their partial information. This feature is apt to be useful in the field biology, where identification markers may be lost, external conditions may cause the premature end of observations, or the observation period may be too brief for all positive events to occur.

Here we consider three examples in which the construction of confidence bands for the difference of two quantile residual life functions allows us to extract useful information in the biomedical and in the ecological context.

Analysis of radio-chemotherapy data We present an analysis of the two-sample data of Stablein and Koutrouvelis (1985) concerning the effects of chemotherapy and chemotherapy plus radiotherapy on the survival times of gastric cancer patients. This example is also analyzed in Hsieh (2001), Kleinbaum (1996), Klein and Moeschberger (1997) and Bagdonavicius et al. (2004). The number of patients is 90. Kaplan and Meier estimators (Kaplan and Meier (1958)) of survival functions pertaining to both treatment groups clearly show a crossing effect phenomenon. The two estimated curves indicate that radiotherapy would initially be detrimental to a patient's survival but becomes beneficial later on. See, for example, Stablein and Koutrouvelis (1985).

We have constructed the 90%-confidence bands for the difference of the median residual life functions for the patients belonging to both groups, see Figure 22. Since there exist censored data, we have considered the median residual life function estimator proposed in Csörgő (1987). From the figure we cannot conclude that one treatment is better than the other over all the support of the random variables.
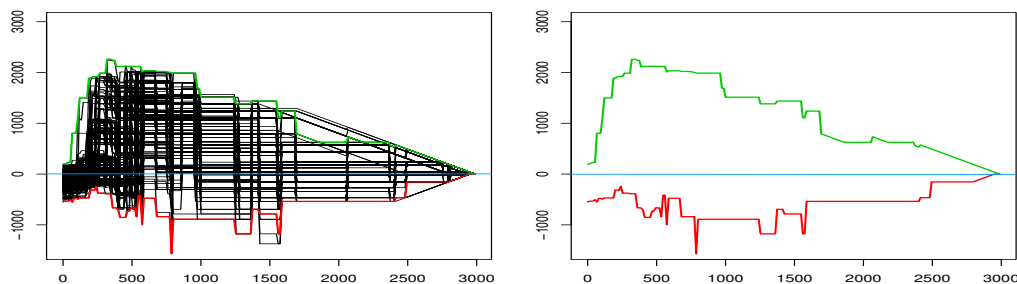

Figure 22: 90%-confidence bands for comparing the two treatments of cancer.

Analysis of ecological data A number of ecological questions can be phrased in terms of 'time until an event occurs'. Events of interest might include the arrival of a migrant

17

or parasite, the display of a particular behavior, the dispersal of a fruit or offspring, the germination of a seed, the abscission of a flower, or the death of an organisms or part of an organism.

Male plants of dioecious species are often more floriferous than female plants, see Lloyd and Webb (1977). This is true of *Clematis ligusticifolia* Nutt., the species we have considered in our example. These data were collected in Matthews-Winters Park, Jefferson County, Colorado, and are available in Muenchow (1986). In this paper it is tested whether males and females are equally attractive to insects against the alternative hypothesis that males are more attractive. The event was defined as the arrival of any flying insect at one of the flowers. He concluded that male flowers were visited at a significantly faster rate than were female flowers after carrying out the Cox-Mantel test.

We have constructed the 90%-confidence bands for the difference of the median residual life functions for both groups of plants, see Figure 23. Again, there exist censored data so we have considered the median residual life function estimator proposed in Csörgő (1987). From the figure we can not conclude that one group of plants is more attractive to insects than the other.
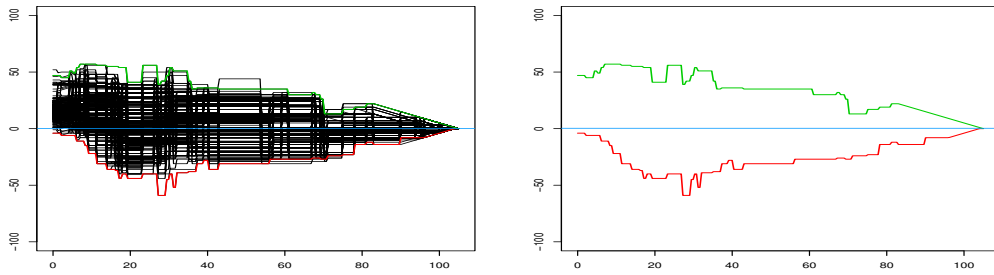

Figure 23: 90%-confidence bands for comparing the two groups of plants.

Analysis of Berger et al. (1988) data The context of this third example involves the influence of different diets on the aging process in rats where research indicates that diet restriction promotes longevity. Yu et al. (1982) studied the effects of a restricted diet on rats versus an *ad libitum* diet, i.e., free eating (see also Witten (1985)).

Let X represent the lifelength in days of rats following *ad libitum* diet and let Y represent the lifelength in days of rats following a restricted diet. Berger et al. (1988) hypothesized

$$H_0 : e_X(t) \leq e_Y(t) \quad \text{for some } t \in [0, 894]$$

versus

$$H_a : e_X(t) > e_Y(t) \quad \text{for some } t \in [0, 894],$$

where e_X and e_Y denote the mean residual life function of X and Y, respectively. At the approximate level $\alpha = 0.1$ the null hypothesis is rejected in favor of H_a.

Figure 24 represents the 90%-confidence bands for the difference of the surviving times of both groups of rats $merl_Y(t) - merl_X(t)$, for all t. The conclusion of the bands is that those

rats following a restricted diet life longer that those following an *ad libitum* diet. Therefore, this analysis confirm the conclusions of Berger et al. (1988) and complete it because they were not able to state that one variable dominates the other for all $t > 894$.
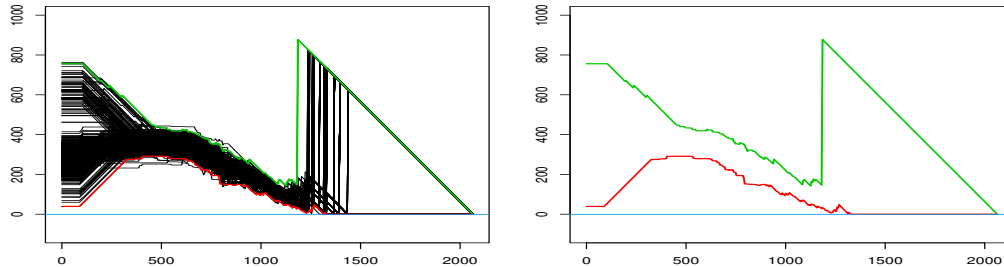

Figure 24: 90%-confidence bands for comparing the two groups of data of Berger et al.

Acknowledgements

References

[1] Arnold, B. C. and Brockett, P. L. (1983). When does the βth quantile residual life function determine the distribution? *Operations Research* **31**, 391–396.

[2] Bagdonavicius, V., Hafdi, M. A. and Nikulin, M. (2004). Analysis of survival data with cross-effect of survival functions *Biostatistics* **5**, 415–425.

[3] Berger, R. L., Boos, D. D. and Guess, F. M. (1988). Tests and confidence sets for comparing two mean residual life functions. *Biometrics* **44**, 103–115.

[4] Cheng, K. F. (1985). Test for equality of failure rates. *Biometrika* **72**, 211–215.

[5] Csörgő, M. (1983). *Quantile processes with statistical applications*, Cambridge University Press.

[6] Csörgő, S. (1987). Estimating quantile residual life under random censorship. *Contributions to stochastics: in honour to the 75th birthday of Walther Eberl, Sr.*, Springer-Verlag.

[7] Cuevas, A., Febrero, M. and Fraiman, R. (2007). Robust estimation and classification for functional data via projection-based depth notions. *Computational Statistics* **22**, 481–496.

[8] Cuevas, A. and Fraiman, R. (2009). On depth measures and dual statistics. A methodology for dealing with general data. *Journal of Multivariate Analysis* **100**, 753–766.

[9] Efron, B. (1981). Censored data and the bootstrap. *Journal of the American Statistical Association* **76**, 312–319.

[10] Fraiman, R. and Meloche, J. (1999). Multivariate *L*-estimation. *Test* **8**, 255–317.

[11] Fraiman, R. and Muniz, G. (2001). Trimmed means for functional data. *Test* **10**, 419–440.

[12] Franco-Pereira, A. M., Lillo, R. E., Romo, J. and Shaked, M. (2010). Percentile residual life orders. *Applied Stochastic Models in Business and Industry*, to appear.

[13] Gupta, R. C. and Langford, E. S. (1984). On the determination of a distribution by its median residual life function: A functional equation. *Journal of Applied Probability* **21**, 120–128.

[14] Haines, A. L. and Singpurwalla, N. D. (1974). Some contributions to the stochastic characterization of wear. *Reliability and Biometry, Statistical Analysis of Lifelength* (edited by F. Proschan and R. J. Serfling), SIAM, Philadelphia, 47–80.

[15] Hsieh, F. (2001). On the heteroscedastic hazards regression models: theory and application. *Journal of the Royal Statistical Society, Series B* **63**, 63–79.

[16] Jeong, J., Jung, S. and Costantino, J. P. (2008) Nonparametric Inference on Median Residual Life Function. *Biometrics* **64**, 157–163.

[17] Joe, H. (1985). Characterizations of life distributions from quantile residual lifetimes. *Annals of the Institute of Statistical Mathematics* **37**, 165–172.

[18] Joe, H. and Proschan, F. (1984a). Percentile residual life functions. *Operations Research* **32**, 668–678.

[19] Joe, H. and Proschan, F. (1984b). Comparison of two life distributions on the basis of their quantile residual life functions. *Canadian Journal of Statistics* **12**, 91–97.

[20] Jung, S., Jeong, J. and Bandos, H. Regression on quantile residual life. *Biometrics* **165**, 1203–1212.

[21] Kaplan, E.L. and Meier, P. (1958). Nonparametric estimation from incomplete observations. *Journal of the American Statistical Association* **53**, 457–481.

[22] Klein, J. P. and Moeschberger, M. L. (1997). *Survival Analysis*, Springer, New York.

[23] Kleinbaum, D. G. (1996). *Survival Analysis; A Serf-Learning Text*, Springer, New York.

[24] Koshevoy, G. and Mosler, K. (1997). Zonoid trimming for multivariate distributions. *The Annals of Statistics* **25**, 1998–2017.

[25] Lillo, R. E. (2005). On the median residual lifetime and its aging properties: A characterization theorem and applications. *Naval Research Logistics* **52**, 370–380.

[26] Lai, C. D. and Xie, M. (2006). *Stochastic Ageing and Dependence for Reliability*, Springer, New York.

[27] Liu R. (1990). On a notion of data depth based on randomsimplices. *The Annals of Statistics* **18**, 405–414.

[28] López-Pintado, S. and Romo, J. (2005). A half-graph depth for functional data. *Working paper* **05-16**.

[29] López-Pintado, S. and Romo, J. (2009). On the concept of depth for functional data. *Journal of the American Statistical Association* **104**, 718–734.

[30] Lloyd D. C. and Webb C. J. (1977). Secondary sex characters in plants. *Botanical Review.* **43**, 177–216.

[31] Mahalanobis, P. C. (1936). On the generalized distance in statistics. *Proceedings of National Academy of Science of India* **12**, 49–55.

[32] Muenchow G. (1986). Ecological use of failure time analysis. *Ecology* **67**, 246–250.

[33] Oja, H. (1983). Descriptive statistics for multivariate distributions. *Statistics and Probability Letters* **1**, 327–332.

[34] Pinder, J. E., Wiener, J. G. and Smith, M. H. (1978). The Weibull distribution: a new method of summarizing survivorship data. *Ecology* **59**, 175–179.

[35] Raja Rao, B., Alhumoud, J. M. and Damaraju, C. V. (2006). Percentile residual life function for a class of life distributions having the 'setting the clock back to zero' property. *International Mathematical Forum* **1**, 1339–1358.

[36] Ramsay, J. O. and Silverman, B. W. (2005). *Functional data analysis*, Springer-Verlag.

[37] Singh, K. (1991). A notion of majority depth. Unpublished document.

[38] Stablein, D. M. and Koutrouvelis, I. A. (1985). A two sample test sensitive to crossing hazards in uncensored and singly censored data. *Biometrics* **41**, 643–652.

[39] Tukey, J. (1975). Mathematics and picturing data. *Proceedings of the 1975 International Congress of Mathematics* **2**, 523–531.

[40] Vardi, Y. and Zhang, C. H. (2000). The multivariate L_1-median and associated data depth. *Proceedings of the National Academy of Science USA* **97**, 1423–1426.

[41] Witten, M. (1985). Reliability theoretic methods and aging: Critical elements, hierarchies, and longevity - Interpreting survival curves. In *Molecular Biology of Aging* A. D. Plackett and A. Hollaender (eds), 345–360. New York: Plenum Press.

[42] Yu, B. P., Masoro, E. J., Murata, I., Bertrand, H. A. and Lynd, F. T. (1982). Lifespan study of SPF Fisher 344 male rats fed *ad libitum* or restricted diets: Longevity, growth, lean body mass and disease. *Journal of Gerontology* **37**, 130–141.

[43] Zuo, Y. (2003). Projection based depth functions and associated medians. *The Annals of Statistics* **31**, 1460–1490.