Evaluation of aging properties by a scale-space inspection of the TTT curve☆

Gámiz, M.L.[1], Nozal-Cañadas, R., Raya-Miranda, R.

University of Granada

Abstract

This template helps you to create a properly formatted LaTeX manuscript.

Keywords: Total-Tim-on-Test Transform, Kernel smoothing, Order statistics, SiZer map
2010 MSC: 90B25, 62N05

1. Introduction

In many applications (e.g. reliability, maintainability, biometry or survival) different forms of aging are of interest giving rise to a well-known non-parametric classification of life distributions (Barlow and Proschan, 1975). While "positive aging" describes the adverse effects of age on the lifetime of components or systems and describes the situation where the residual life tends to decrease in some probabilistic sense with increasing age, "no aging" means that age has no effect on residual life of an item, and "negative agin" means that the item experiences certain improvement in performance or reliability with age, see Lay *et al.* (2006). We are concerned with "positive aging" however it has to be mentioned that the classification according to different types of "positive aging" has a parallel in terms of "negative aging".

In a technical contex, aging is usually understood as a gradual decrease in performance over time, so it can be expected that the aging of a component or system increases the probability that it will fail.

Email address: mgamiz@ugr.es (Gámiz, M.L.)
[1]Corresponding author

If the lifetime of a mechanism is a random variable X, a natural measure of aging could be the probability of surviving above $t+x$, assuming that the mechanism has an age of at least t, this idea leads to the concept of residual life that will be formally defined later in this paper. Another possible measure of aging is formulated in terms of the hazard function, which measures the instantaneous propensity to the failure of a mechanism as a function of its age. An inspection of the shape of these functions is very useful to evaluate the performance of a system in the future which is very important for making decisions in maintainability and inventory theory, for example.

The most relevant classes of life distributions are denoted in the literature as IFR, IFRA, NBU, NBUE and DMRL (with duals) and can be found in Barlow and Proschan (1975), Kochar and Singh (1988) and Cao and Wang (1991), among others. Although, the mathematical description of aging is done under three broad categories based on hazard functions, residual life functions and survival functions, in this paper we will focus on the Total-Time-on-Test curve (TTT) to describe a useful and novel graphical tool to characterize some important types of aging. The TTT plot was introduced by Barlow and Campo (1975) as a tool for analysing failure data.

Parts of the paper:

- Definition and properties of the Total-Time-on-Test (TTT) transform
- Local-polynomial fit of the TTT curve and its derivatives
- SiZer map for studying convexity properties of a curve.
- SiZer map for other aging properties
- Hypothesis testing

2. Total-Time-on-Test transform: Definition and relevant properties

Some important classes of life distributions (see Bergman and Klefsjo, 1983) can be characterized in terms of the Total-Time-on-Test curve. We recall some of the most well-known classes but first we need the following definition. We assume that X is a random lifetime with absolutely continuous distribution function.

Definition **2.1. Residual life**
Let X a lifetime with survival function S, for any fixed $t > 0$ we denote X_t the residual lifetime from time t, which represent the additional lifetime from t. The conditional survival function is obtained as $S_t(x) = S(t+x)/S(t)$, for all $x > 0$. The mean residual time is then defined as $\mu_t = \int_0^{+\infty} \frac{S(t+x)}{S(t)} dx$.

Definition **2.2. Aging classes**
Let X be a lifetime with distribution function F, survival function S and finite mean μ.

1. *Increasing (decreasing) failure rate, $IFR(DFR)$:*
   X is IFR(DFR) if and only if its conditional survival function is increasing (decreasing). That is, for all $x > 0$ fixed, and for all $t_1 < t_2$ we have that $S_{t_1}(x) \leq (\geq) S_{t_2}(x)$. This condition can be also stablished in terms of the hazard rate, saying that X is IFR(DFR) its hazard function is increasing (resp. decreasing).
2. *Upside-down Failure Rate, UFR: X is UFR if and only if its hazard function is initially increasing to a pick after declining abruptly till stabilize.*
3. *Bathtub shaped Failure Rate, BFR: X is (BFR) if and only if its hazard rate curve is divided into three regions: decreasing hazard rate region, constant hazard rate region, and increasing hazard rate region.*
4. *Decreasing (increasing) mean residual life, $DMRL(IMRL)$:*
   X is DMRL(IMRL) if and only if μ_t is decreasing (increasing) for all $t > 0$.
5. *New better (worse) than used in expectation, $NBUE(NWUE)$:*
   X is NBUE (NWUE) if and only if $\mu_t \leq (\geq) \mu$, for all $t > 0$; where $\mu_t = \int_0^{+\infty} \frac{S(t+x)}{S(t)} dx$, denotes the expected residual life of a mechanism with distribution F and at the age t.

2.1. Definition of TTT curve

Total time on test plots provide a useful graphical method for tentative identification of lifetime models. This concept is very important in applications in reliability analysis. When several units are tested for studying their life lengths some of the units would fail while others may survive the test period. The sum of all observed and incomplete life lengths is the total time on test statistics. The plot of this statistic versus time is called the total-time-on-test plot. As the number of units on test tends to infinity the limit of this statistic is called the total time on test transform (TTT). This concept was introduced by [**?**] and later studied by [4]. Model identification is based on properties of the TTT transform.

Definition **2.3.** ***Total-Time-on-Test statistic***
Suppose n *items under test and successive failures are observed at* $0 = X_{(0)} \leq X_{(1)} \leq X_{(2)} \leq \cdots \leq X_{(n)}$, *with* $X_{(i)}$ *the* i*-th order statistic from a lifetime random variable* X *with absolutely continuous distribution function* F. *Then the total time on test statistic during the interval* $(0, t)$ *is defined as*

$$\tau(t) = \sum_{i=1}^{r} X_{(i)} + (n-r)\, t,$$

provided that $X_{(r-1)} < t \leq X_{(r)}$.

For comparison purposes, usually the statistic is scaled to the interval $[0, 1]$ by means of the transformation $\tau\left(X_{(r)}\right) / \tau\left(X_{(n)}\right)$, for $r = 1, 2, \ldots, n$.

Based on the sample order statistics we can construct the empirical distribution function, that is $F_n(t) = r/n$, for $X_{(r)} \leq t < X_{(r+1)}$, for $r = 1, 2, \ldots, n-1$; with $F_n(t) = 0$, for $t < X_{(1)}$; and, with $F_n(t) = 1$, for $t \geq X_{(n)}$. We can define the following function

$$F_n^{-1}(p) = \inf\left\{x \geq 0 : F_n(x) > p\right\},$$

it can be checked, Nair *et al.* (2010), that

$$\int_0^{F_n^{-1}\left(\frac{r}{n}\right)} (1 - F_n(t))\, dt = \frac{\tau\left(X_{(r)}\right)}{n},$$

and also that

$$\lim_{n\to\infty} \lim_{\frac{r}{n}\to p} \int_0^{F_n^{-1}\left(\frac{r}{n}\right)} (1 - F_n(t))\, dt = \int_0^{F^{-1}(p)} (1 - F(t))\, dt,$$

uniformly in $p \in (0, 1)$.

Definition **2.4.** ***Total Time on Test transform***
Let X *be a (non-negative) random variable with absolutely continuous cumulative distribution function* F. *The total time on test transform of* X *is defined as*

$$\varphi(p) = \int_0^{Q(p)} (1 - F(t))\, dt \tag{1}$$

where we denote $Q(p) = F^{-1}(p)$, for $p \in [0,1]$, the corresponding quantile function.

When $E[X] < \infty$, this expectation can be obtained as $\mu = \int_0^{Q(1)} S(t)dt$, where we denote $S(t) = 1 - F(t)$, the survival function. Then we define the ***scaled*** TTT transform as $\varphi(p)/\mu$, for $0 \leq p \leq 1$, which is scale invariant. We keep notation $\varphi(p)$ for the scaled TTT transform, and we assume that $\mu = 1$, since it does not imply any loss of generality.

2.2. Aging properties based on the TTT transform

The scaled TTT transform can be used to characterize different aging properties, see Barlow and Proschan, (1975) and Bergman and Lindqvist (1983). For F the exponential distribution, it can be checked that $\varphi(p) = p$, for $0 \leq p \leq 1$. As mentioned above, based on a sample $X_1, X_2, \ldots, X_n$ one can construct the TTT plot, which will be closer to the scaled TTT curve as n tends to $+\infty$. We can thus use the TTT plot as a tool for model selection, in the sense that when, for example, the TTT plot produces a cloud of points around the diagonal of the square unit, we can admit that the distribution of the underlying lifetime F is exponential, that is with constant hazard rate. Other hazard shapes can be recognized from an inspection of the TTT plot. A convex TTT curve corresponds with a decreasing hazard (DFR); a concave TTT curve corresponds with an increasing hazard (IFR). When the TTT plot describes a trajectory first convex then concave it indicates a bathtub hazard shape and when it is concave then convex, it indicates a unimodal hazard shape.

In summary, to determine the type of aging represented by X we can study the shape of the TTT curve. To do it we compute the first and second derivatives of the TTT transform. Using expression (1) can be obtained as

$$\begin{aligned} \varphi'(p) &= \frac{d}{dp}\left\{\int_0^{Q(p)} S(x)dx\right\} = S\left(Q(p)\right)Q'(p) = (1-p)Q'(p); \text{ and,} \\ \varphi''(p) &= -Q'(p) + (1-p)Q''(p). \end{aligned}$$

where $Q'(p) = dQ(p)/dp$ and $Q''(p) = d^2Q(p)/d^2p$.

Example **2.1.** *Exponential distribution Let X be a random variable with distribution $Exp(\lambda)$.....*

Let X be a lifetime with distribution function F and with finite mean μ. From the theorems given in Klefsjo (1980) the above aging conditions can be expressed in terms of the TTT-curve as follows

Proposition **2.1.** *TTT-curve characterization of aging*

1. *F is $IFR(DFR)$ if and only if $\varphi''(u) \leq (\geq) 0$, for all $0 < u < 1$;*
2. *F is $DMRL(IMRL)$ if and only if $\varphi'(u) \leq (\geq) \frac{1-\varphi(u)}{1-u}$;*
3. *F is $NBUE(NWUE)$ if and only if $\varphi(u)/u \leq (\geq) 1$, for all $0 < u < 1$.*

3. Local polynomial estimation of the TTT curve and its first and second derivatives

In this section we suggest a local-polynomial estimator for the TTT-transform directly formulated from an empirical estimation of the TTT-curve, that is, we do not need to estimate the quantiles to get an estimator of the TTT-transform.

3.1. Least-squares estimation of the Total-Time-on-Test transform

Let $X_1, X_2, \cdots, X_n$ be an independent and identically distributed random sample drawn form an absolutely continuous distribution function F with density f. Let $X_{(1)}, X_{(2)}, \cdots, X_{(n)}$ denote the corresponding order statistics. Let us denote $p_i = \frac{i}{n}$ for $i = 1, 2, \ldots, n$. An empirical (*naive*) estimator of the TTT-curve $\widehat{\varphi}_n$, can be defined as follows

$$\widehat{\varphi}_n\left(p_i\right) = \sum_{j=1}^{i} \left(1 - \frac{j-1}{n}\right) \left(X_{(j)} - X_{(j-1)}\right), \tag{2}$$

for $i = 1, 2, \ldots, n$, with $\widehat{\varphi}_n(0) = 0$. We can observe that $\widehat{\varphi}_n(1) = \bar{X}$, the mean sample statistics. Since the properties of the curve are not affected by scale changes, we can confine the curve to be defined in the interval $[0, 1]$ by first normalizing the data. That is, we work with the sample $\{X_i/\bar{X}; i = 1, \ldots, n\}$. In the following, without loss of generality, we assume that $\bar{X} = 1$.

Under a local-polynomial approach we consider that, for each estimation point $p_0 \in (0, 1)$, the TTT-transform $\varphi(p_0)$ is locally (in a neighborhood of p_0) approximated by a mth-degree polynomial function in the sense that for all $p \in (p_0 - h, p_0 + h)$ we have that $\varphi(p) = \theta_0 + \theta_1(p - p_0) + \theta_2(p - p_0)^2 + \ldots + \theta_m(p - p_0)^m$, for an appropriate bandwidth h.

The parameters of the model can be interpreted respectively as $\theta_0 = \varphi(p_0)$; $\theta_1 = \varphi'(p_0)$; and, in general, $\theta_k = \frac{\varphi^{(k)}(p_0)}{k!}$, for $k = 1, 2, \ldots, m$. The approximation above is valid locally if we assume certain smoothness conditions on the quantile function, in the sense of derivability.

In particular, the three first coefficients provide the following estimates: $\widehat{\varphi}(p_0) = \widehat{\theta}$, $\widehat{\varphi'}(p_0) = \widehat{\theta}_1$, and $\widehat{\varphi''}(p_0) = 2\widehat{\theta}_2$, respectively. To this goal, we formulate the following least squares problem

$$\begin{aligned}&\left(\widehat{\theta}_0, \widehat{\theta}_1, \ldots, \widehat{\theta}_m\right)^{\intercal} = \\ &\mathop{\arg\min}_{(\theta_0,\theta_1,\ldots,\theta_m)^{\intercal}} \sum_{i=1}^{n} \left\{\widehat{\varphi}_n(p_i) - \theta_0 - \theta_1(p_i - p_0) - \ldots - \theta_m(p_i - p_0)^m\right\}^2 K_h(p_i - p_0),\end{aligned} \tag{3}$$

where $K_h(\cdot) = \frac{1}{h}K(\frac{\cdot}{h})$, with K a symmetric kernel function, and h the bandwidth parameter that controls the size of the window around the estimation point p_0 where the polynomial approximation is valid. After differentiating in equation (3), for $m = 2$ with respect to θ_j ($j = 0, 1, 2$), we obtain a system of linear equations that can be written in matrix form

$$\begin{pmatrix} a_0 & a_1 & \cdots & a_m \\ a_1 & a_2 & \cdots & a_{m+1} \\ \vdots & \vdots & \ddots & \vdots \\ a_m & a_{m+1} & \cdots & a_{2m} \end{pmatrix} \begin{pmatrix} \theta_0 \\ \theta_1 \\ \vdots \\ \theta_m \end{pmatrix} = \begin{pmatrix} A_0 \\ A_1 \\ \vdots \\ A_m \end{pmatrix};$$

where we have denoted

$$A_r(p_0) = \sum_{i=1}^{n} \widehat{\varphi}_n(p_i)(p_i - p_0)^r K_h(p_i - p_0), \quad r = 0,\ 1, \ldots m;$$

and

$$a_l(p_0) = \sum_{i=1}^{n} (p_i - p_0)^l K_h(p_i - p_0), \quad l = 0,\ 1,\ 2, \ldots, 2m.$$

3.2. Local-quadratic estimator

In particular we can fit the data locally by using a 2-degree polynomial. We denote this estimator $\widehat{\varphi}_{2,h}(\cdot)$. Then, we set the least squares problem of (3) for $m = 2$ and define

$$A_r(p_0) = \sum_{i=1}^{n} \widehat{\varphi}_n(p_i)(p_i - p_0)^r K_h(p_i - p_0), \quad r = 0,\ 1,\ 2;$$

and

$$a_l(p_0) = \sum_{i=1}^{n} (p_i - p_0)^l K_h(p_i - p_0), \quad l = 0,\ 1,\ 2,\ 3,\ 4.$$

After differentiating in equation (3), for $m = 2$ with respect to θ_j ($j = 0, 1, 2$), we obtain a system of linear equations that can be written in matrix form

$$\begin{pmatrix} A_0 \\ A_1 \\ A_2 \end{pmatrix} = \begin{pmatrix} a_0 & a_1 & a_2 \\ a_1 & a_2 & a_3 \\ a_2 & a_3 & a_4 \end{pmatrix} \begin{pmatrix} \theta_0 \\ \theta_1 \\ \theta_2 \end{pmatrix}. \tag{4}$$

Using Cramer's rule, the solution can be then expressed as

$$\widehat{\theta}_0 = \frac{\begin{vmatrix} A_0 & a_1 & a_2 \\ A_1 & a_2 & a_3 \\ A_2 & a_3 & a_4 \end{vmatrix}}{\begin{vmatrix} a_0 & a_1 & a_2 \\ a_1 & a_2 & a_3 \\ a_2 & a_3 & a_4 \end{vmatrix}} = \widehat{\varphi}_h(p_0), \qquad \widehat{\theta}_1 = \frac{\begin{vmatrix} a_0 & A_0 & a_2 \\ a_1 & A_1 & a_3 \\ a_2 & A_2 & a_4 \end{vmatrix}}{\begin{vmatrix} a_0 & a_1 & a_2 \\ a_1 & a_2 & a_3 \\ a_2 & a_3 & a_4 \end{vmatrix}} = \widehat{\varphi'}_h(p_0),$$

$$\widehat{\theta}_2 = \frac{\begin{vmatrix} a_0 & a_1 & A_0 \\ a_1 & a_2 & A_1 \\ a_2 & a_3 & A_2 \end{vmatrix}}{\begin{vmatrix} a_0 & a_1 & a_2 \\ a_1 & a_2 & a_3 \\ a_2 & a_3 & a_4 \end{vmatrix}} = 2\ \widehat{\varphi''}_h(p_0),$$

where we denote $|\mathbf{A}|$, the determinant of matrix $\mathbf{A}$. We then get the estimation of the curve φ and its derivatives at a given p_0, based on the quadratic fit as follows.

1. The TTT-curve:

$$\widehat{\varphi}_h(p_0) = \widehat{\theta}_0 = \sum_{i=1}^{n} \bar{K}_{0,h}\,(p_i - p_0)\,\widehat{\varphi}_n(p_i), \tag{5}$$

where

$$\bar{K}_{0,h}\,(p_i - p_0) = \frac{(a_1 a_3 - a_2^2)\,(p_i - p_0)^2 + (a_2 a_3 - a_1 a_4)\,(p_i - p_0) + (a_2 a_4 - a_3^2)}{a_0 a_2 a_4 + 2 a_1 a_2 a_3 - a_2^3 - a_0 a_3^2 - a_1^2 a_4} K_h\,(p_i - p_0)\,.$$

2. The first derivative:

$$\widehat{\varphi'}_h(p_0) = \widehat{\theta}_1 = \sum_{i=1}^{n} \bar{K}_{1,h}\left(p_i - p_0\right)\widehat{\varphi}_n(p_i), \tag{6}$$

where

$$\bar{K}_{1,h}\left(p_i - p_0\right) = \frac{\left(a_1 a_2 - a_0 a_3\right)\left(p_i - p_0\right)^2 + \left(a_0 a_4 - a_2^2\right)\left(p_i - p_0\right) + \left(a_2 a_3 - a_1 a_4\right)}{a_0 a_2 a_4 + 2a_1 a_2 a_3 - a_2^3 - a_0 a_3^2 - a_1^2 a_4} K_h\left(p_i - p_0\right).$$

3. The second derivative:

$$\widehat{\varphi''}_h(p_0) = 2\widehat{\theta}_2(p_0) = 2\sum_{i=1}^{n} \bar{K}_{2,h}\left(p_i - p_0\right)\widehat{\varphi}_n(p_i) \tag{7}$$

where we denote

$$\bar{K}_{2,h}\left(p_i - p_0\right) = \frac{\left(a_0 a_2 - a_1^2\right)\left(p_i - p_0\right)^2 + \left(a_1 a_2 - a_0 a_3\right)\left(p_i - p_0\right) + \left(a_1 a_3 - a_2^2\right)}{a_0 a_2 a_4 + 2a_1 a_2 a_3 - a_2^3 - a_0 a_3^2 - a_1^2 a_4} K_h\left(p_i - p_0\right).$$

3.3. Local-cubic estimator

After solving the equations in (4) with $m = 3$, and taking $\widehat{\varphi}_n(\cdot)$ the empirical estimator given in (2), we can write the estimators of φ and its first derivatives similar to the previous section as follows.

1. The local-cubic TTT-curve estimate:

$$\widetilde{\varphi}_h(p_0) = \widetilde{\theta}_0(p_0) = \sum_{i=1}^{n} \widetilde{K}_{0,h}\left(p_i - p_0\right)\widehat{\varphi}_n\left(p_i\right) \tag{8}$$

where

$$\widetilde{K}_{0,h}\left(p_i - p_0\right) = \frac{\Delta_{00} - \Delta_{01}\left(p_i - p_0\right) + \Delta_{02}\left(p_i - p_0\right)^2 - \Delta_{03}\left(p_i - p_0\right)^3}{\Delta} K_h\left(p_i - p_0\right)$$

and with

$$\Delta_{00} = \begin{vmatrix} a_2 & a_3 & a_4 \\ a_3 & a_4 & a_5 \\ a_4 & a_5 & a_6 \end{vmatrix};\ \Delta_{01} = \begin{vmatrix} a_1 & a_2 & a_3 \\ a_3 & a_4 & a_5 \\ a_4 & a_5 & a_6 \end{vmatrix};\ \Delta_{02} = \begin{vmatrix} a_1 & a_2 & a_3 \\ a_2 & a_3 & a_4 \\ a_4 & a_5 & a_6 \end{vmatrix};$$

$$\Delta_{03} = \begin{vmatrix} a_1 & a_2 & a_3 \\ a_2 & a_3 & a_4 \\ a_3 & a_4 & a_5 \end{vmatrix};\quad \Delta = \begin{vmatrix} a_0 & a_1 & a_2 & a_3 \\ a_1 & a_2 & a_3 & a_4 \\ a_2 & a_3 & a_4 & a_5 \\ a_3 & a_4 & a_5 & a_6 \end{vmatrix}.$$

2. The first derivative:

$$\widetilde{\varphi'}_h(p_0) = \widetilde{\theta}_1(p_0) = \sum_{i=1}^{n} \widetilde{K}_{1,h}\left(p_i - p_0\right) \widehat{\varphi}_n\left(p_i\right), \tag{9}$$

where

$$\widetilde{K}_{1,h} = \frac{-\Delta_{1,0} + \Delta_{1,1}\left(p_i - p_0\right) - \Delta_{1,2}\left(p_i - p_0\right)^2 + \Delta_{1,3}\left(p_i - p_0\right)^3}{\Delta} K_h\left(p_i - p_0\right),$$

and with

$$\Delta_{10} = \begin{vmatrix} a_1 & a_3 & a_4 \\ a_2 & a_4 & a_5 \\ a_3 & a_5 & a_6 \end{vmatrix}; \ \Delta_{11} = \begin{vmatrix} a_0 & a_2 & a_3 \\ a_2 & a_4 & a_5 \\ a_3 & a_5 & a_6 \end{vmatrix}; \ \Delta_{12} = \begin{vmatrix} a_0 & a_2 & a_3 \\ a_1 & a_3 & a_4 \\ a_3 & a_5 & a_6 \end{vmatrix};$$

$$\Delta_{13} = \begin{vmatrix} a_0 & a_2 & a_3 \\ a_1 & a_3 & a_4 \\ a_2 & a_4 & a_5 \end{vmatrix}.$$

Finally the second derivative, based on local-cubic approximation.

3. The second derivative:

$$\widetilde{\varphi''}_h(p_0) = \sum_{i=1}^{n} \widetilde{K}_{2,h}\left(p_i - p_0\right) \widehat{\varphi}_n\left(p_i\right), \tag{10}$$

where

$$\widetilde{K}_{2,h}\left(p_i - p_0\right) = 2\widetilde{\theta}_2(p_0) = 2\frac{\Delta_{2,0} - \Delta_{2,1}\left(p_i - p_0\right) + \Delta_{2,2}\left(p_i - p_0\right)^2 - \Delta_{2,3}\left(p_i - p_0\right)^3}{\Delta} K_h\left(p_i - p_0\right.$$

and with

$$\Delta_{20} = \begin{vmatrix} a_1 & a_2 & a_4 \\ a_2 & a_3 & a_5 \\ a_3 & a_4 & a_6 \end{vmatrix}; \ \Delta_{21} = \begin{vmatrix} a_0 & a_1 & a_3 \\ a_2 & a_3 & a_5 \\ a_3 & a_4 & a_6 \end{vmatrix}; \ \Delta_{22} = \begin{vmatrix} a_0 & a_1 & a_3 \\ a_1 & a_2 & a_4 \\ a_3 & a_4 & a_6 \end{vmatrix};$$

$$\Delta_{23} = \begin{vmatrix} a_0 & a_1 & a_3 \\ a_1 & a_2 & a_4 \\ a_2 & a_3 & a_5 \end{vmatrix};$$

3.4. Statistical properties of the estimator

To derive statistical properties we write the estimators given above in terms of the corresponding L-estimator. That is, we re-write expressions in (5)-(7) and (8)-(10) each as a linear combination of the order statistics $X_{(1)} \leq X_{(2)} \leq \cdots \leq X_{(n)}$. Then we will use the results in Hutson and Ernst (2000) to obtain exact analytic expression for the bootstrap moments, in particular we will need the mean, and variance. We give the details for the local-quadratic case, the results for the local-cubic estimator can be derived directly.

The local-quadratic estimator

We start from the empirical estimator $\widehat{\varphi}_n$ detailed in (2), which can be expressed as

$$\widehat{\varphi}_n(p_i) = \sum_{j=1}^{i} \omega_{i,j} X_{(j)},$$

where the weights $\omega_{i,j}$, are given by

$$\omega_{i,j} = \begin{cases} \frac{1}{n}, & j = 1, 2, \ldots, i-1; \\ \frac{n-(i-1)}{n}, & j = i. \end{cases}$$

We can arrange these weights in matrix of the form

$$\mathbf{W} = \begin{pmatrix} 1 & 0 & 0 & 0 & \cdots & 0 \\ \frac{1}{n} & \frac{n-1}{n} & 0 & 0 & \cdots & 0 \\ \frac{1}{n} & \frac{1}{n} & \frac{n-2}{n} & 0 & \cdots & 0 \\ \frac{1}{n} & \frac{1}{n} & \frac{1}{n} & \frac{n-3}{n} & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \frac{1}{n} & \frac{1}{n} & \frac{1}{n} & \cdots & \frac{1}{n} & \frac{1}{n} \end{pmatrix}$$

The TTT-curve estimator

From (5), we have that

$$\begin{aligned} \widehat{\varphi}_h(p_0) &= \sum_{i=1}^{n} \bar{K}_{0,h}(p_i - p_0) \sum_{j=1}^{i} \omega_{i,j} X_{(j)} \\ &= \sum_{i=1}^{n} \sum_{j=1}^{i} \bar{K}_{0,h}(p_i - p_0) \omega_{i,j} X_{(j)} \end{aligned}$$

re-ordering the terms in this expression we get

$$\widehat{\varphi}_h(p_0) = \sum_{j=1}^{n} \sum_{i=j}^{n} \bar{K}_{0,h}(p_i - p_0)\omega_{i,j} X_{(j)}, \tag{11}$$

Using matrix notation we can write the above expression (11) as

$$\widehat{\varphi}_h(p_0) = \bar{\mathbf{K}}_{0,h}(p_0)^{\intercal} \cdot \mathbf{W} \cdot \mathbf{X}_{(\cdot)}, \tag{12}$$

where we have denoted the vector of weights

$$\bar{\mathbf{K}}_{0,h}(p_0) = \left(\bar{K}_{0,h}(p_1 - p_0), \bar{K}_{0,h}(p_2 - p_0), \ldots, \bar{K}_{0,h}(p_n - p_0)\right)^{\intercal},$$

and $\mathbf{X}_{(\cdot)} = \left(X_{(1)}, \ldots, X_{(n)}\right)^{\intercal}$ is the ordered sample.

We conclude that the expression (12) is an L-estimator. Let us define the following characteristics of the sample. The vector of means of the order statistics

$$\boldsymbol{\mu} = \left(\mu_{(1)}, \mu_{(2)}, \ldots, \mu_{(n)}\right)^{\intercal}, \tag{13}$$

with $\mu_{(i)} = E\left[X_{(i)}\right]$, for $i = 1, 2, \ldots, n$, and the covariance matrix

$$\boldsymbol{\Sigma} = \begin{pmatrix} \sigma^2_{(1)} & \sigma_{(12)} & \cdots & \sigma_{(1n)} \\ \sigma_{(12)} & \sigma^2_{(2)} & \cdots & \sigma_{(2n)} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{(1n)} & \sigma_{(2n)} & \cdots & \sigma^2_{(n)} \end{pmatrix} \tag{14}$$

being$\sigma^2_{(i)} = \mathrm{Var}(X_{(i)})$ and $\sigma_{(ij)} = \mathrm{Cov}\left(X_{(i)}, X_{(j)}\right) = E\left[\left(X_{(i)} - \mu_{(i)}\right) \cdot \left(X_{(j)} - \mu_{(j)}\right)\right]$, for $i, j = 1, 2, \ldots, n$, $i < j$.

Then, it is easy to check that the variance of the estimator of the TTT-curve can be expressed as

$$\mathrm{Var}\left(\widehat{\varphi}_h(p_0)\right) = \bar{\mathbf{K}}_{0,h}(p_0)^{\intercal} \cdot \mathbf{W} \cdot \boldsymbol{\Sigma} \cdot \mathbf{W}^{\intercal} \cdot \bar{\mathbf{K}}_{0,h}(p_0) \tag{15}$$

The first derivative of the TTT-curve

Similar to (15) we can obtain the corresponding expression for the first derivative of the TTT-curve as

$$\widehat{\varphi'}_h(p_0) = \sum_{j=1}^{n} \sum_{i=j}^{n} \bar{K}_{1,h}(p_i - p_0)\omega_{i,j} X_{(j)} = \bar{\mathbf{K}}_{1,h}(p_0)^{\intercal} \cdot \mathbf{W} \cdot \mathbf{X}_{(\cdot)},$$

where we have defined the corresponding vector of weights

$$\bar{\mathbf{K}}_{1,h}(p_0) = \left(\bar{K}_{1,h}(p_1 - p_0), \bar{K}_{1,h}(p_2 - p_0), \ldots, \bar{K}_{1,h}(p_n - p_0)\right)^{\intercal},$$

the variance is

$$\operatorname{Var}\left(\widehat{\varphi'}_h(p_0)\right) = \bar{\mathbf{K}}_{1,h}(p_0)^{\intercal} \cdot \mathbf{W} \cdot \boldsymbol{\Sigma} \cdot \mathbf{W}^{\intercal} \cdot \bar{\mathbf{K}}_{1,h}(p_0) \tag{16}$$

The second derivative of the TTT-curve

Finally for the local-quadratic estimator of the second derivative of the TTT-curve we write

$$\widehat{\varphi''}_h(p_0) = 2\sum_{j=1}^{n}\sum_{i=j}^{n} \bar{K}_{2,h}(p_i - p_0)\omega_{i,j}X_{(j)} = 2\bar{\mathbf{K}}_{2,h}(p_0)^{\intercal} \cdot \mathbf{W} \cdot \mathbf{X}_{(\cdot)},$$

where we have denoted

$$\bar{\mathbf{K}}_{2,h}(p_0) = \left(\bar{K}_{2,h}(p_1 - p_0), \bar{K}_{2,h}(p_2 - p_0), \ldots, \bar{K}_{2,h}(p_n - p_0)\right)^{\intercal},$$

then, we obtain the variance as

$$\operatorname{Var}\left(\widehat{\varphi''}_h(p_0)\right) = 4\bar{\mathbf{K}}_{2,h}(p_0)^{\intercal} \cdot \mathbf{W} \cdot \boldsymbol{\Sigma} \cdot \mathbf{W}^{\intercal} \cdot \bar{\mathbf{K}}_{2,h}(p_0) \tag{17}$$

The local-cubic estimator

For the local-cubic estimator, we proceed similarly.

$$\begin{aligned} \widetilde{\varphi}_h(p_0) &= \sum_{i=1}^{n} \widetilde{K}_{0,h}(p_i - p_0) \sum_{j=1}^{i} \omega_{i,j}X_{(j)} \\ &= \sum_{i=1}^{n}\sum_{j=1}^{i} \widetilde{K}_{0,h}(p_i - p_0)\omega_{i,j}X_{(j)}, \end{aligned}$$

which is also an L-estimator. Let us denote

$$\widetilde{\mathbf{K}}_{0,h}(p_0) = \left(\widetilde{K}_{0,h}(p_1 - p_0), \widetilde{K}_{0,h}(p_2 - p_0), \ldots, \widetilde{K}_{0,h}(p_n - p_0)\right)^{\intercal},$$

so we can write the variance of the estimator of the TTT-curve using the local-cubic approach as

$$\mathrm{Var}\left(\widetilde{\varphi}_h(p_0)\right) = \left(\widetilde{\mathbf{K}}_{0,h}(p_0)\mathbf{W}\right)^{\intercal} \cdot \mathbf{\Sigma} \cdot \left(\widetilde{\mathbf{K}}_{0,h}(p_0)\mathbf{W}\right) \tag{18}$$

for each $p_0 \in (0,1)$. Equally we can obtain the variance for the first and second derivatives of the local-cubic estimate.

3.5. Moments of order statistics

In this section we use bootstrap techniques to estimate the first two moments of the order statistics. That is, we estimate the elements of vector $\boldsymbol{\mu}$, given in (13), as well as the covariance matrix $\mathbf{\Sigma}$ given in (14). Let us denote the corresponding bootstrap estimators $\widehat{\boldsymbol{\mu}}$ and $\widehat{\mathbf{\Sigma}}$ respectively. We consider on the one hand exact analytic expressions for these bootstrap mean and covariance matrix, and also we obtain the corresponding estimators based on Montecarlo resampling.

Exact amalytic expressions

Let $X_1, X_2, \ldots, X_n$ be a sample of independent random variables with a common absolutely continuous distribution F, and let $X_{(1)} \leq X_{(2)} \leq \cdots X_{(n)}$ be the order statistics.

First, we define mth non-centered moment of the rth order statistics as (see Arnold, Balakrishnan and Nagaraja, 2008)

$$\mu_{(r)}^{(m)} = E\left[X_{(r)}^m\right] = \frac{n!}{(r-1)!(n-r)!}\int_0^1 (Q(u))^m u^{r-1}(1-u)^{n-r}du, \tag{19}$$

for $1 \leq r \leq n$, and with $Q(u) = F^{-1}(u)$ the quantile function. We focus on $m = 1, 2$, because we are only interested in the mean and the variance, and as usual we compute $\sigma_{(r)}^2 = \mu_{(r)}^{(2)} - \left(\mu_{(r)}\right)^2$.

We can replace $Q(u)$ by $\widehat{Q}(u)$, and considering that $\widehat{Q}(u) = X_{(j)}$ for $(j-1)/n < u \leq j/n$ we obtain

$$\widehat{\mu}_{(r)}^{(m)} = \sum_{j=1}^{n}\int_{(j-1)/n}^{j/n} (Q(u))^m f_B(u)du\ dv = \sum_{j=1}^{n} X_{(j)}^m \int_{(j-1)/n}^{j/n} f_B(u)du, \tag{20}$$

where we denote $f_B(u) = \frac{n!}{(r-1)!(n-r)!} u^{r-1}(1-u)^{n-r}$, the density function of a Beta distribution with parameters $a_1 = r$ and $a_2 = n - r + 1$. Then we conclude that the corresponding moment can be easily computed using the incomplete beta function, that is

$$\widehat{\mu}_{(r)}^{(m)} = \sum_{j=1}^{n} X_{(j)}^{m} \left[F_B(\frac{j}{n}; a_1 = r, a_2 = n-r+1) - F_B(\frac{j-1}{n}; a_1 = r, a_2 = n-r+1) \right], \tag{21}$$

where $F_B(x; a_1, a_2) = \frac{n!}{(r-1)!(n-r)!} \int_0^x u^{a_1-1}(1-u)^{a_2-1} du$.

For $1 \leq r < s \leq n$, the (1,1)-order moment can be defined as

$$\mu_{(rs)} = E\left[X_{(r)}X_{(s)}\right] =_n C_{rs} \int_0^1 \int_0^u Q(u)Q(v)u^{r-1}(v-u)^{s-r-1}(1-u)^{n-s} dvdu; \tag{22}$$

where ${}_nC_{rs} = \frac{n!}{(r-1)!(s-r-1)!(n-s)!}$

The double integral in (22) can be computed as

$$\begin{aligned} \widehat{\mu}_{(rs)} &= {}_nC_{rs} \sum_{j=2}^{n} \sum_{i=1}^{j-1} \int_{(j-1)/n}^{j/n} \int_{(i-1)/n}^{i/n} \widehat{Q}(u)\widehat{Q}(v)u^{r-1}(v-u)^{s-r-1}(1-u)^{n-s} dvdu + \\ &+ \sum_{j=1}^{n} \int_{(j-1)/n}^{j/n} \int_{(j-1)/n}^{u} \widehat{Q}(u)\widehat{Q}(v)u^{r-1}(v-u)^{s-r-1}(1-u)^{n-s} dvdu \end{aligned} \tag{23}$$

Again we replace the quantile function that appears in the integrand by its empirical version, so we get

$$\begin{aligned} \widehat{\mu}_{(rs)} &= \sum_{j=2}^{n} \sum_{i=1}^{j-1} X_{(i)}X_{(j)} \int_{(j-1)/n}^{j/n} \int_{(i-1)/n}^{i/n} {}_nC_{rs} u^{r-1}(v-u)^{s-r-1}(1-u)^{n-s} dvdu + \\ &+ \sum_{j=1}^{n} X_{(j)}^2 \int_{(j-1)/n}^{j/n} \int_{(j-1)/n}^{u} {}_nC_{rs} u^{r-1}(v-u)^{s-r-1}(1-u)^{n-s} dvdu \end{aligned} \tag{24}$$

The function inside the integrals can be seen as the density function of a Dirichlet distribution, which is defined as

$$f_D(u, v) = \frac{\Gamma(a_1 + a_2 + a_3)}{\Gamma(a_1)\Gamma(a_2)\Gamma(a_3)} u^{a_1-1} v^{a_2-1} (1-u-v)^{a_3-1} \tag{25}$$

where $0 < u, v < 1$, and $0 < u + v < 1$.

Estimation of the moments by Monte carlo simulation

Let $\{X_1, X_2, \ldots X_n\}$ denote a sample of independent lifetimes identically distributed as X . The aim is to estimate the mean and moments of order 2 defined in the previous section by using resampling techniques. We propose Monte carlo simulation as explained in the following algorithm.

Algorithm 1. Bootstrapped moments of the order statistics

Step 1. Draw with replacement a total of n items from the set $\{1, 2, \ldots, n\}$, that is: $i_1, i_2, \ldots, i_n$;

Step 2. For each i_j, take the corresponding X_{i_j}, for $j = 1, \ldots, n$. Denote $X^*_{(1)} \leq X^*_{(2)} \leq \ldots \leq X^*_{(n)}$ the resulting bootstrap order statistics sequence.

Step 3. Repeat Step 2 up to M times. Construct the $n \times M$-dimensional matrix of the form

$$\mathbf{X}^* = \begin{pmatrix} X^1_{(1)} & X^2_{(1)} & \ldots & X^M_{(1)} \\ X^1_{(2)} & X^2_{(2)} & \ldots & X^M_{(2)} \\ \cdots & \cdots & \ddots & \vdots \\ \ldots & \ldots & \ddots & \vdots \\ X^1_{(n)} & X^2_{(n)} & \ldots & X^M_{(n)} \end{pmatrix}$$

Step 4. Define the vector of bootstrap means $\widehat{\boldsymbol{\mu}}^*$, with components $\widehat{\mu}^*_j = (1/M) \sum_{m=1}^{M} X^m_{(j)}$, for $j = 1, 2, \ldots, n$; and the bootstrap covariance matrix $\widehat{\boldsymbol{\Sigma}} = (1/M)\,(\mathbf{X}^* - \widehat{\boldsymbol{\mu}}^*)^\intercal\,(\mathbf{X}^* - \widehat{\boldsymbol{\mu}}^*)$.

4. Development of SiZer map for evaluating aging properties

4.1. Description of the SiZer Map

SiZer, a graphical tool introduced by Chaudhuri and Marron (1999, 2000), has proven to be a powerful methodolody for conducting exploratory data analysis. At a bump, there is a zero crossing of the derivative: all estimated slope with different bandwidths to its left are significantly increasing while all estimated slopes to its right are significantly decreasing. "SiZer" relies on three plots, the so-called "family plot", which is the representation of a family of non-parametric smoothers of the target function, indexed by the bandwidth parameter, the gradient SiZer map, which displays the scale and space inference about the first derivative. For each bandwidth, which corresponds to the scale, and each value in the support, which gives localization, a confidence interval for the first derivative is calculated and the signs are displayed on the map using a color code. Here we use a black-and-white version of "SiZer". Considering n_h scales and p_0 localizations, each pixel in the $(p_0 \times n_h)$ map is coded as white if zero is greater than the upper confidence bound, indicating significant decrease; black if zero is less than the lower confidence bound, indicating significant increase; gray if zero is within the confidence limits (no significant increase or decrease); and dark gray indicating regions where the data are too sparse to infer significance. The structure in the data is highlighted by gray changes and a change from black to white means a significant bump. The last, the SiCon map, which displays the scale and space inference about the second derivative. For this map, the black-and-white squeme is black for negative curvature (concave), white for positive curvatuve (convex), gray for zero curvature, and dark gray indicating regions where the dara are too sparse.

Poner un ejemplo con datos simulados, e incluirle la verdadera curva para poder comparar. Ejemplo que presenta el family plot, primera derivada y segunda derivada. Especificar aqu que en los siguientes ejemplo no se mostrar la primera derivada porque siempre ser azul (o los ponemos siempre aunque sean siempre iguales para verificar.)

(Cambiar esta interpretacin por la que corresponda al ejemplo)The top panel of Figure 1 shows a family of smoothers of hazard rates, in this case local linear kernel hazard estimates, for various bandwidth parameter values. The thick line corresponds to the true hazard function. This reveals the characteristics of the underlying hazard rate. Larger bandwidths produce oversmoothed curves showing the trend, while small bandwidths provide too detailed a description of the hazard rate. Figure 1 shows the inference about the first derivative of the hazard rate. The vertical axis has different smoothing levels, among which the user has to choose the most significant one. The horizontal axis shows the values where the hazard rate is evaluated (localizations). Figure 1 shows that the hazard rate increases significantly at the beginning (black color up to about the point 0.1). There is then a zero crossing signaled by the gray color (between points 0.1 and 0.3), indicating that the hazard rate is constant. The hazard rate then decreases to the point 0.5, as the white color indicates. This reveals a bump around 0.2. A second bump may occur around the value 0.8 (black goes up to 0.8, indicating a significant increase, followed by gray (zero crossing of the derivative), and ending with white (significant decrease)). This interpretation is in concordance with the true underlying hazard. The beta functions defining the true model are the density function B(t; 2,6) which has a bump at 0, and B(t; 6,2) which has a bump around 1. The inference with "SiZer" manages to detect these two features of the true hazard.

4.2. "SiZer" for evaluating aging trends

In this section, we take the general ideas behind SiZer methodology (Chaudhuri and Marron 1999, 2000) and adapt it to be able to inspect a specific aging property that underlies the lifetime being analyzed.

The IFR(DFR) family of lifetime distributions

In this case we examine the sign of the second derivative of the TTT curve. Then we have calculated

Confidence interval for $\varphi''(p)$

For $0 \leq p \leq 1$,

$$\left[\widehat{\varphi}''_h(p) - q_{1-\frac{\alpha}{2}}\sqrt{\mathrm{Var}\left(\widehat{\varphi}''_h(p)\right)},\ \widehat{\varphi}''_h(p) + q_{1-\frac{\alpha}{2}}\sqrt{\mathrm{Var}\left(\widehat{\varphi}''_h(p)\right)}\right]$$

More general properties

SiZer map techniques are developed for other classes of life distributions more general than the IFR (DFR) class.

- The NBUE (NWUE) family of lifetime distributions.

  In this case we examine the sign of the following transformation of the TTT curve: $g(p) = \varphi(p)/p$.

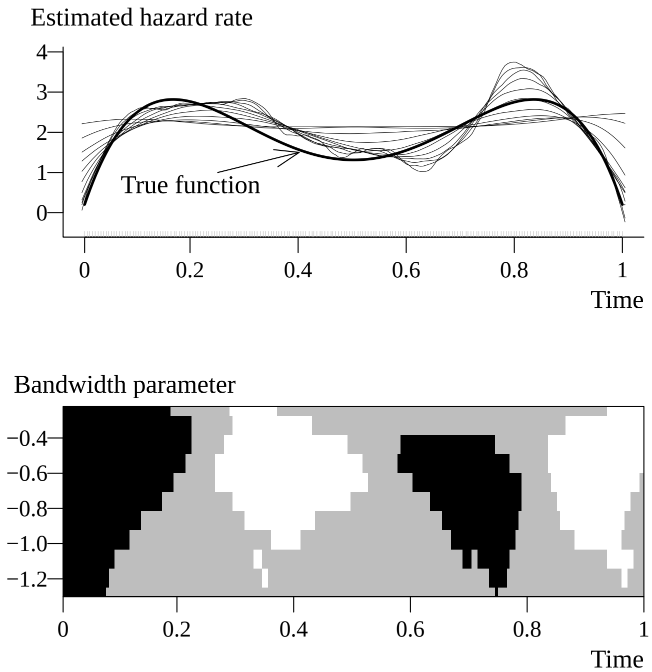

Figure 1: “SiZer” with simulated data. From left to right, the change black, white, black, then white indicates three changes in trend, revealing two significant bumps.

Confidence interval for $g(p)$

For $0 \leq p \leq 1$,

$$\left[\frac{\widehat{\varphi}_h(p) - q_{1-\frac{\alpha}{2}}\sqrt{\mathrm{Var}\left(\widehat{\varphi}_h(p)\right)}}{p},\ \frac{\widehat{\varphi}_h(p) + q_{1-\frac{\alpha}{2}}\sqrt{\mathrm{Var}\left(\widehat{\varphi}_h(p)\right)}}{p}\right]$$

- The DMRL (IMRL) family of lifetime distributions

  Este caso no estoy segura de incluirlo*****.....

4.3. Hypothesis testing based on SiZer Map

We are firstly interested in checking whether the data are exponentially distributed. In such a case the true TTT curve coincides with the diagonal of the unit square. Let X be a random lifetime, we want to test the null hypothesis

$$\mathrm{H}_0 : X \text{ follows } \mathrm{Exp}(\lambda) \text{ distribution, for some} \lambda > 0,$$

against general alternatives based on a sample $X_1, X_2, \ldots, X_n$ of independent copies of X. To solve the problem we propose a new procedure based on the TTT-transform and that uses the SiZer tool. In other words, we set the following hypothesis testing problem:

$$\begin{aligned} &\mathrm{H}_0 : \varphi''_X(p) = 0, \text{for all } p \in (0,1) \\ &\mathrm{H}_1 : \varphi''_X(p) \neq 0, \text{for some } p \in (0,1) \end{aligned} \tag{26}$$

We can rewrite the hypotheses in (26) in SiZer map language, then the null hypothesis is equivalent to assert that the underlying map to the true distribution is completely purple. We will call this one the *true* map. We will decide to reject the null hypothesis in case that the SiZer map based on the *empirical* map (referred as displayes) a percentage of non purple pixels above a pre-specified level. As usual we will refer to this level the type I error probability, and denote it as α. This value is commonly taken as $\alpha = 0.05$.

We summarize the steps of our proposal in the following algorithm.

Algorithm 2. Testing exponentiality vs. aging trend

Step 1. Compute the sample mean $\bar{X}$ and define $\bar{X}_i = X_i/\bar{X}$, for $i = 1, 2, \ldots, n$;

Step 2. Generate M bootstrap samples as explained in Algorithm 1, and construct the corresponding $\mathbf{X}^*$, matrix.

Step 3. Construct M empirical SiZer maps for the second derivative as explained in Section.... One for each bootstrap sample of each column of matrix $\mathbf{X}^*$;

Step 4. Compare each *empirical* map with the *true* map pixel by pixel, and count the total number of pixels where the color in the generated *empirical* map is not the same as in the *true* map. Define a binary function δ taking value 1 when the corresponding empirical map reports one or more than one non-purple pixels;

Step 5. Define the bootstrap p-value as $\sum_{m=1}^{M} \delta_m / M$

Step 6. Reject the null hypothesis when p- value lower than α

5. Simulations

6. Real data examples

7. Concluding remarks

The aim of this paper has been to .....

References

[1]

[2] Barlow, R.E. and Campo, R. Total time on test processes and applications to failure data analysis. In Reliability and Fault Tree Analysis, Edited by: Barlow, R.E., Fussell, J. and Singpurwalla, N.D. (1975) 451481

[3] Barlow and Doksum (1972) Isotonic test for convex ordering. Procedings of the $6th$ Berkeley Symposium on Mathematical Statistics and Probability, Vol. 1,pp. 293–323, University of California Press

[4] Barlow, R. E. and F. Proschan (1975). Statistical theory of reliability and life testing. Holt, Rinehart and Winston.

[5] Hutson, A.D., and Ernst, M. (2000) The exact bootstrap mean and variance of an L-estimator, J.R. Statist. Soc B. Vol. 62, Part 1, pp. 89–94.

[6] Nair *et al.* (2010)