Ann. Inst. Statist. Math.
Vol. 47, No. 3, 483–491 (1995)

CHANGE POINT ESTIMATION IN NON-MONOTONIC AGING MODELS

MURARI MITRA[1] AND SUJIT K. BASU[2]

[1] *Division of Theoretical Statistics and Mathematics, Indian Statistical Institute, 203, B.T. Road, Calcutta 700 035, India*
[2] *Indian Institute of Management, Joka, Diamond Harbour Road, Post Office Box No. 16757, Calcutta 700 027, India*

(Received March 14, 1994; revised December 5, 1994)

Abstract. The problem of estimating change points in various non-monotonic aging models is considered. A general methodology for consistent estimation of the change point is developed and applied to non-monotonic aging models based on the hazard rate function as well as on the mean residual life function.

Key words and phrases: Non-monotonic aging, NWBUE, IDMRL, BFR life distributions, change points.

1. Introduction

The various nonparametric classes of life distributions common in reliability theory, deal with positive (or negative) aging according as the residual lifetime tends to decrease (or increase), in some probabilistic sense, with increasing age of the component/individual under consideration. This gives rise to different monotonic aging classes such as IFR, DMRL, NBU, NBUE and their duals depending on whether the aging pattern is positive or negative. However, in many practical situations, the effect of age is initially beneficial (a 'burn-in' phase where negative aging takes place), but after a certain period, it is adverse indicating a 'wear-out' phase where aging is positive. This kind of non-monotonic aging is typically modelled using life distributions displaying bathtub failure rates (BFR). Guess *et al.* (1986) and Mitra and Basu (1994) have introduced the IDMRL and NWBUE families of distributions respectively, to study the above phenomenon through the mean residual life (MRL) function.

For a life distribution function (d.f.) $F(\cdot)$, the MRL function at age $x \geq 0$ is defined as $e_F(x) := E(X - x \mid X > x) = (1/\bar{F}(x)) \int_x^\infty \bar{F}(t)dt$ whenever $\bar{F}(x) := 1 - F(x) > 0$ and 0 otherwise.

DEFINITION 1.1. (Guess *et al.* (1986)) A life distribution F with support $[0, \infty)$ is said to be an *Increasing initially, then Decreasing Mean Residual Life*

(IDMRL) distribution if there exists a $x_0 \geq 0$ such that $e_F(x)$ is *non-decreasing* on $[0, x_0)$ and *non-increasing* on (x_0, ∞); in such a case we say that 'F is IDMRL(x_0)'.

Example 1.1. Consider the life distribution characterised by the survival function

$$\bar{F}(x) = \begin{cases} 4/(2+x)^2 & \text{if } 0 \leq x < 1 \\ (4x/9)\exp[-(x^2-1)/6] & \text{if } 1 \leq x < \infty. \end{cases}$$

The corresponding MRL function is as follows:

$$e_F(x) = \begin{cases} 2+x & \text{if } 0 \leq x < 1 \\ 3/x & \text{if } x \geq 1. \end{cases}$$

It is evident that F is IDMRL(x_0) with unique change point $x_0 = 1$.

DEFINITION 1.2. A life d.f. F having support $[0, \infty)$ (and finite mean μ) is said to be *New Worse then Better than Used in Expectation* (NWBUE) if there exists a point $x_0 \geq 0$ such that

$$e_F(x) \begin{cases} \geq (\leq) e_F(0), & \text{for } x < x_0, \\ \leq (\geq) e_F(0), & \text{for } x \geq x_0; \end{cases}$$

in such a case, we write 'F is NWBUE(x_0)'.

It can be readily seen that the distribution function introduced in Example 1.1 is also NWBUE(x_0) with unique change point $x_0 = 3/2$. It can be shown that $\{\text{IDMRL}\} \subseteq \{\text{NWBUE}\}$ (Proposition 2.3 of Mitra and Basu (1994)). That $\{\text{NWBUE}\}$ is a strictly larger class of life d.f.s is borne out by the following example.

Example 1.2. Consider the life distribution having the following survival function:

$$\bar{F}(x) = \begin{cases} \{\theta/(\theta+x)\}^2 & \text{if } 0 \leq x < \alpha \\ \{\theta/(\theta+\alpha)\}^2 & \text{if } \alpha \leq x < \beta \\ \{\theta/(\theta+\alpha)\}^2[(\theta+2\alpha-\beta)/(\theta+2\alpha-2\beta+x)]^2 & \text{if } \beta \leq x < \gamma \\ \dfrac{\theta^2}{(\theta+\alpha)^2}\dfrac{(\theta+2\alpha-\beta)^2}{(\theta+2\alpha-2\beta+\gamma)^2}(x/\gamma) & \\ \quad \cdot \exp\left[-\dfrac{1}{2}\dfrac{(x^2-\gamma^2)}{\gamma(\theta+2\alpha-2\beta+\gamma)}\right] & \text{if } x \geq \gamma \end{cases}$$

where the parameters α, β, γ and θ are such that $0 < \alpha \leq \beta \leq \gamma$, $2\alpha > \beta$ and $\theta > 0$. The corresponding MRL function then works out as

$$e_F(x) = \begin{cases} \theta + x & \text{if } 0 \leq x < \alpha \\ \theta + 2\alpha - x & \text{if } \alpha \leq x < \beta \\ \theta + 2\alpha - 2\beta + x & \text{if } \beta \leq x < \gamma \\ \gamma(\theta+2\alpha-2\beta+\gamma)/x & \text{if } x \geq \gamma. \end{cases}$$

It is easy to observe that the life distribution F is NWBUE with unique change point $x_0 = \gamma(\theta + 2\alpha - 2\beta + \gamma)/\theta$; however, F is *not* IDMRL whenever $\alpha < \beta < \gamma$.

For the purpose of the next definition, we require the notion of a failure rate function. Whenever F is absolutely continuous with a probability density function (pdf) $f(\cdot)$, we define $r(\cdot) = f(\cdot)/\bar{F}(\cdot)$ to be its failure rate function.

DEFINITION 1.3. A life distribution F having support $[0, \infty)$ is said to be a *Bathtub Failure Rate* (BFR) distribution if there exists a $x_0 \geq 0$ such that $r(t)$ is *non-increasing* on $[0, x_0)$ and *non-decreasing* on $[x_0, \infty)$; in such a case, we write 'F is BFR(x_0)'.

We shall refer to such an x_0 (which need not be unique) as a change point of a d.f. F in the BFR sense and write F is BFR(x_0). Similarly, the change points in the two previous definitions can be referred to as change points in the IDMRL and NWBUE sense respectively.

As has been shown in Mitra and Basu (1994), $\{\text{BFR}\} \subseteq \{\text{NWBUE}\}$ and accordingly, the BFR distributions mentioned in Rajarshi and Rajarshi (1988) are also NWBUE. It may be mentioned at this juncture that the NWBUE family is strictly larger than the BFR family as has been demonstrated in Example 2.2 of Mitra and Basu (1994).

In the next section, we outline a procedure for estimating such change points and demonstrate that our method applies in all the above mentioned classes. Earlier efforts in estimating change points were limited to *specific* life distributions only, rather than the issue being addressed to general nonparametric classes of life distributions at large. For example, Nguyen *et al.* (1984) and Yao (1986) considered the life distribution having failure rate function

$$r(t) = \alpha\chi(0 \leq t \leq \tau) + \beta\chi(t > \tau)$$

where $\chi(A)$ is the indicator function of the set A and $\alpha, \beta, \tau \in \mathbb{R}$. The following general model characterized by the hazard function

$$r(t) = \sum_{k=1}^{m} \alpha_k \chi(\tau_{k-1} \leq t < \tau_k)$$

with $0 = \tau_0 < \tau_1 < \cdots < \tau_m = \infty$ and $\alpha_k \geq 0 \forall k$, was subsequently considered by Pham and Nguyen (1990). They used techniques of maximum likelihood estimation (the 'pseudo-maximum likelihood' approach) and established strong consistency of the suggested estimators. The application of bootstrap methods in the case $m = 2$ in this model has been studied by Pham and Nguyen (1993). Basu *et al.* (1988) treated the 'truncated bathtub model' specified by the rate function

$$r(t) = \lambda(t)\chi(0 \leq t \leq \tau) + \lambda_0\chi(t > \tau),$$

where $\lambda(t)$ is a decreasing positive function and λ_0 is a positive constant. As opposed to the above parametric and semiparametric models, here we propose to consider the problem of estimating the change point in purely nonparametric setups. An earlier effort in this direction has been by Kulasekera and Lal Saxena (1991) who considered life distributions displaying bathtub failure rates having a unique change point and proposed a consistent estimator of the same. The methodology described in the next section when applied to the BFR case, seems to work under less restrictive assumptions.

2. Change point estimation

We consider BFR, IDMRL and NWBUE life distributions having *unique* change points and suggest a unified approach for estimating these. Estimation of change points is relevant particularly in the context of maintenance policies, since, as is natural, one would hardly think of preventively replacing a component having such a life distribution before the 'threshold' (unknown) age of x_0 is achieved.

In a general setup, our problem can be formulated as follows:

Given a random sample, $X_1, X_2, \ldots, X_n$ of size n from an unknown life d.f. F, where F is NM(x_0), $x_0 < \infty$, (NM is the abbreviation for non-monotonic and in subsequent discussions it would stand for either BFR or IDMRL or NWBUE as the case may be), how to estimate the unknown change point x_0?

We shall make the following two assumptions:

(A1) The finite change point x_0 of the d.f. F, whenever positive, is the *unique positive minimizer/maximizer* of a suitable non-negative transform $h_F(\cdot)$ of F; otherwise, it is the *unique minimizer/maximizer*.

(A2) We have an upper bound for the unknown change point—call it B, i.e., $x_0 \le B < \infty$, $F(B) < 1$.

Note that (A2) is quite a weak assumption because, in most practical situations, an idea about B can be formed on the basis of some prior experience concerning the phenomenon under consideration.

While estimating the change point of a IDMRL distribution F, we shall take $h_F(x)$ to be $e_F(x)$; likewise, in the BFR case, $h_F(x)$ is taken to be $r_F(x)$. Consequently, in the context of both these cases, the assumption (A1) is equivalent to

(A1*) The finite change point x_0 of the d.f. F is unique.

However, this equivalence fails in the NWBUE case where we choose $h_F(x) \equiv |e_F(x) - \mu_F|$ with μ_F being the mean of F. This is so because here, under (A1*), $h_F(x) = 0$ may have multiple solutions, exactly one of which is the change point.

The unified approach for change point estimation in either of the three cases consists in identifying x_0 as the *unique* minimizer (or maximizer, as the case may be) of a specific non-negative transform $h_F(x)$ of F. We then estimate $h_F(x)$ by $h_{F_n^*}(x)$, where F_n^* is a *suitable* estimator of F and is based on a random sample of size n from it. Depending on the case, let Λ_n be the set of minimizers or maximizers (not exceeding B) of $h_{F_n^*}(x)$. We then estimate x_0 by x_{0_n}, where $x_{0_n} := \inf \Lambda_n$ whenever Λ_n is non-empty; x_{0_n} is defined appropriately otherwise. The choice of F_n^* would depend on the specific problem being tackled as will be evident from the three cases dealing with the estimation of the change point when F is (i) NWBUE(x_0), (ii) IDMRL(x_0) and (iii) BFR(x_0).

(i) *The* NWBUE(x_0) *Case.* Suppose F is NWBUE(x_0) with finite mean μ_F. Consider the non-negative transform defined by

$$h_F(x) := |e_F(x) - \mu_F|. \tag{2.1}$$

Note that, in this context, condition (A1) indicates that whenever $x_0 = 0$, it is the *unique* minimizer of $h_F(x)$ and it is the *unique positive* minimizer otherwise. Let F_n be the empirical c.d.f. based on the random sample $X_1, X_2, \ldots, X_n$ and $X_{(1)} \le X_{(2)} \le \cdots \le X_{(n)}$ be the corresponding order statistics. In this case, we estimate $h_F(x)$ by

$$h_{F_n}(x) := |e_{F_n}(x) - \mu_{F_n}|, \tag{2.2}$$

where $e_{F_n}(x)$ is given by

$$e_{F_n}(x) = \begin{cases} \dfrac{1}{\bar{F}_n(x)} \displaystyle\int_x^\infty \bar{F}_n(v)dv & \text{if } X_{(n)} > x, \\ 0 & \text{otherwise,} \end{cases}$$

and μ_{F_n} is the sample mean $\bar{X}_n$. Note that $e_{F_n}(x)$ is nothing but the usual life table estimate of the life expectancy at age x. Also, it is easy to see that

$$e_{F_n}(x) = \begin{cases} \dfrac{\sum_{j=1}^n (X_j - x) I(X_j - x)}{\sum_{j=1}^n I(X_j - x)} & \text{if } X_{(n)} > x, \\ 0 & \text{otherwise,} \end{cases} \tag{2.3}$$

where $I(a) = 1$ or 0 according as $a > 0$ or $a \le 0$. It follows from Yang (1978) that as $n \to \infty$,

$$\sup_{0 \le x \le b} |e_{F_n}(x) - e_F(x)| \to 0 \quad \text{a.s.} \tag{2.4}$$

for every $b > 0$ satisfying $F(b) < 1$. Further, it is clear from (2.3) that $e_{F_n}(x)$ is a right continuous function having finite left limits and is piecewise linearly decreasing in the intervals $[0, X_{(1)}), [X_{(1)}, X_{(2)}), \ldots, [X_{(n-1)}, X_{(n)})$. Also, note that $e_{F_n}(x)$ has jumps at $x = X_{(i)}$, $i = 1, 2, \ldots, n-1$. These jumps are of *positive* magnitude, since the averages $(n - k + 1)^{-1} \sum_{j=k}^n X_{(j)}$ increase with k. Now, define

$$\Lambda_n := \{0 < x \le B : |e_{F_n}(x) - \bar{X}_n| \text{ is minimum}\}. \tag{2.5}$$

Observe that $\lim_{x\downarrow 0}|e_{F_n}(x)-\bar{X}_n| = 0$ with probability 1; hence, if Λ_n is non-empty, the minimum value of $|e_{F_n}(x)-\bar{X}_n|$ has to be zero; we thus have,

$$\Lambda_n = \{0 < x \leq B : e_{F_n}(x) = \bar{X}_n\}.$$

Then we have the following useful lemma:

LEMMA 2.1. *If $x_0 > 0$ $(=0)$, there exists an integer $n_0 \geq 1$ such that for all $n \geq n_0$, Λ_n is non-empty (empty) with probability* 1.

PROOF. Consider the case $x_0 > 0$. Let $0 < x_1 < x_0 < x_2 < B$. As F is NWBUE(x_0), and x_0 is the unique change point of F,

$$\begin{aligned} &e_F(x_1) - \mu_F > 0, \\ &e_F(x_2) - \mu_F < 0. \end{aligned} \tag{2.6}$$

Using (2.4) and Kolmogorov's SLLN, we note that

$$e_{F_n}(x) - \bar{X}_n \to e_F(x) - \mu_F \quad \text{a.s.} \quad \text{as} \quad n \to \infty.$$

It then follows from (2.6) that there exists $n_0 \geq 1$, sufficiently large, such that for all $n \geq n_0$, $e_{F_n}(x_1) - \bar{X}_n > 0$ and $e_{F_n}(x_2) - \bar{X}_n < 0$ hold with probability 1.

The above inequalities, together with the fact that $e_{F_n}(\cdot)$ has jumps that can only be *positive* in magnitude, guarantee the existence of a solution to the equation $e_{F_n}(x) = \bar{X}_n$ with probability 1, for all sufficiently large n.

If $x_0 = 0$, then via (2.4) and SLLN, we observe that for all $x > 0$, $e_{F_n}(x) - \bar{X}_n > 0$ a.s. $\forall$ large n, since F is then NWBUE with unique change point zero. □

In view of the above lemma, the observation $\Lambda_n = \emptyset$ in a given instance seems to suggest that F is NBUE, i.e. NWBUE with $x_0 = 0$. We are thus motivated to estimate x_0 by 0 in such a situation. On the other hand, when Λ_n is non-empty, we shall estimate x_0 by $x_{0_n} := \inf \Lambda_n$; as a matter of fact, we could estimate x_0 by any member of Λ_n since in the long run, intuitively, we expect Λ_n to be a singleton set by virtue of (2.4) and our assumption (A1).

Remark 2.1. Lemma 2.1 implies that $x_{0_n} = 0$ for all large n with probability 1 whenever $x_0 = 0$.

The following theorem justifies the use of the estimator proposed above.

THEOREM 2.1. *The estimator x_{0_n} is strongly consistent for x_0.*

PROOF. In case $x_0 = 0$, a much stronger conclusion holds in view of Remark 2.1; in fact, beyond a certain stage, x_{0_n} becomes zero identically. Fix any $\omega \in \Omega$, where $(\Omega, \mathcal{F}, P)$ is the probability space on which the X_i's are defined. By definition, $\{x_{0_n}\}$ is a bounded sequence and as such, it has a convergent subsequence. Let $\{x_{0_{n_k}}\}$ be any convergent subsequence of $\{x_{0_n}\}$ and suppose that

$$x_{0_{n_k}} \to x_0^* \quad \text{as} \quad k \to \infty. \tag{2.7}$$

At this stage, we note that the graph of $e_{F_n}(x)$ consists of n linear segments, each having negative slope. Each such segment extends between two successive order statistics. Moreover, $e_{F_n}(0) = \bar{X}_n$ and $e_{F_n}(x)$ is linearly decreasing between 0 and $X_{(1)}$. Thus it is clear that Λ_n can have at most $(n-1)$ elements so that $\inf \Lambda_n = \min \Lambda_n$ and as such $x_{0_{n_k}} \in \Lambda_{n_k}$. Therefore,

$$0 \leq |e_{F_{0_{n_k}}}(x_{0_{n_k}}) - \bar{X}_{n_k}| \leq |e_{F_{0_{n_k}}}(x_0) - \bar{X}_{n_k}|. \tag{2.8}$$

Taking limits as $k \to \infty$ and using (2.4) and the SLLN, we have, with probability 1,

$$0 \leq |e_F(x_0^*) - \mu_F| \leq |e_F(x_0) - \mu_F| = 0.$$

The last equality holds as x_0 solves $e_F(x) = \mu_F$. Thus, with probability 1, $x_0^* = x_0$, by the uniqueness of the change point. Hence, any convergent subsequence of $\{x_{0_n}\}$ converges a.s. to x_0. This completes the proof. □

(ii) *The* IDMRL(x_0) *Case.* Here, we take $h_F(x) \equiv e_F(x)$ so that we can identify x_0 as the unique maximizer of $e_F(x)$ and estimate $e_F(x)$ by $e_{F_n}(x)$ as in Case (i). Define

$$\Lambda_n := \{0 < x \leq B : e_{F_n}(x) \text{ is maximum}\}.$$

It follows from the graph of $e_{F_n}(\cdot)$ that Λ_n is empty if and only if $e_{F_n}(x) - \bar{X}_n < 0$ a.s. $\forall x > 0$; likewise if $e_{F_n}(x) - \bar{X}_n \geq 0$ a.s. for some $x > 0$, then Λ_n will comprise exclusively of one or more of the order statistics. Now (2.4) and the IDMRL property of F imply that for all sufficiently large n, Λ_n is non-empty (empty) with probability 1 whenever $x_0 > 0$ $(= 0)$. We define

$$x_{0_n} = \begin{cases} 0 & \text{if } \Lambda_n = \emptyset \\ \inf \Lambda_n = \min \Lambda_n & \text{if } \Lambda_n \neq \emptyset. \end{cases}$$

It is to be noted that the estimate of the change point would either be zero (if Λ_n is empty) or one of the order statistics (if Λ_n is non-empty).

The consistency of x_{0_n} can be proved very easily along lines similar to those in Case (i).

(iii) *The* BFR(x_0) *Case.* We finally discuss the estimation of the change point x_0 of a BFR distribution F for which the failure rate function $r_F(x)$ is well-defined. For this purpose, we take $h_F(x) \equiv r_F(x)$; under (A1*), we note that x_0 is the *unique* minimizer of $h_F(x)$.

We assume the p.d.f. $f(\cdot)$ of F to be uniformly continuous and let $f_n(x)$ be a continuous kernel estimate of $f(\cdot)$. Accordingly, under suitable assumptions on the kernel function, we have, by Theorem A of Silverman (1978),

$$\sup\{|f_n(x) - f(x)| : x \geq 0\} \to 0 \quad \text{a.s.} \quad \text{as} \quad n \to \infty. \tag{2.9}$$

We propose to estimate $r_F(x)$ by

$$r_n(x) := \frac{f_n(x)}{\bar{F}_n(x)}, \tag{2.10}$$

where

$$\bar{F}_n(x) = 1 - n/(n+1)\int_0^x f_n(y)dy,$$

$f_n(\cdot)$ being the above kernel estimate of $f(\cdot)$. The factor $n/(n+1)$ is introduced in the definition of $\bar{F}_n(\cdot)$ simply to ensure that $r_n(\cdot)$ in (2.9) is well-defined, as $\int_0^\infty f_n(x)dx \leq 1$. It is simple to deduce that

$$|r_n(x) - r(x)| \to 0 \quad \text{a.s.} \quad \text{as} \quad n \to \infty \tag{2.11}$$

uniformly, over any bounded interval.

Define

$$\Lambda_n := \{0 \leq x \leq B : r_n(x) \text{ is minimum}\}. \tag{2.12}$$

Since, for each fixed sample point, $r_n(x)$ is a continuous function of x, it attains its bounds over the compact set $[0, B]$. Thus Λ_n is non-empty, and we set

$$x_{0_n} := \inf \Lambda_n = \min \Lambda_n. \tag{2.13}$$

The last equality holds in view of the continuity of $r_n(\cdot)$.

The estimator proposed in (2.12) seems more natural and intuitively appealing compared to the one given by Kulasekera and Lal Saxena (1991), which looks complicated, besides being computationally rather involved. Moreover, they needed a number of assumptions on the failure rate as well as the kernel function to establish the strong consistency of their estimator. But, the estimator in (2.12) is strongly consistent under much less restrictive conditions, as can be seen in Theorem 2.2 below.

For the purpose of Theorem 2.2 below, we assume the conditions of Theorem A of Silverman (1978), so that (2.9) follows.

THEOREM 2.2. *The estimator x_{0_n} defined in* (2.12) *is strongly consistent for* x_0.

PROOF. Let $\{x_{0_{n_k}}\}$ be a convergent subsequence of the bounded sequence $\{x_{0_n}\}$ and let

$$x_{0_{n_k}} \to x_0^*.$$

By definition of $x_{0_{n_k}}$,

$$r_{n_k}(x_{0_{n_k}}) \leq r_{n_k}(x_0).$$

Taking limits as $k \to \infty$ and using (2.10), we get,

$$r(x_0^*) \leq r(x_0),$$

which completes the proof because of the uniqueness of x_0. $\square$

References

Basu, A. P., Ghosh, J. K. and Joshi, S. N. (1988). On estimating change point in a failure rate, *Statistical Decision Theory and Related Topics IV, No. 2*, 239–252, Springer, New York.

Guess, F., Hollander, M. and Proschan, F. (1986). Testing exponentiality versus a trend change in mean residual life, *Ann. Statist.*, **14**, 1388–1398.

Kulasekera, K. B. and Lal Saxena, K. M. (1991). Estimation of change point in failure rate models, *J. Statist. Plann. Inference*, **29**, 111–124.

Mitra, M. and Basu, S. K. (1994). On a nonparametric family of life distributions and its dual, *J. Statist. Plann. Inference*, **39**, 385–397.

Nguyen, H. T., Rogers, G. S. and Walker, E. A. (1984). Estimation in change point hazard rate models, *Biometrika*, **71**, 299–304.

Pham, T. D. and Nguyen, H. T. (1990). Strong consistency of the maximum likelihood estimators in the change-point hazard rate model, *Statistics*, **21**, No. 2, 203–216.

Pham, T. D. and Nguyen, H. T. (1993). Bootstrapping the change-point of a hazard rate, *Ann. Inst. Statist. Math.*, **45**, No. 2, 331–340.

Rajarshi, M. B. and Rajarshi, S. M. (1988). Bathtub distributions—a review, *Comm. Statist. Theory Methods*, **17**, 2597–2622.

Silverman, B. W. (1978). Strong and weak uniform consistency of the kernel estimate of a density and its derivatives, *Ann. Statist.*, **6**, 177–184.

Yang, G. L. (1978). Estimation of a biometric function, *Ann. Statist.*, **6**, 112–116.

Yao, Y. C. (1986). Maximum likelihood estimation in hazard rate models with a change point, *Comm. Statist. Theory Methods*, **15**, 2455–2466.