Testing failure data for evidence of aging

Yu Hayakawa, Sue Paul and Tony Vignaux
School of Mathematical and Computing Sciences
Victoria University of Wellington
P.O. Box 600, Wellington, New Zealand
yu.hayakawa@vuw.ac.nz, sue.paul@vuw.ac.nz, tony.vignaux@vuw.ac.nz

Abstract

This paper presents an application of a Bayesian nonparametric method to hypothesis testing between exponential and increasing failure rate (IFR) distributions. The weighted gamma process is selected as a prior on the space of non-decreasing failure rates. Monte Carlo simulations of the weighted Chinese restaurant process provide an approximation to the posterior probability of each hypothesis.

1 Introduction

In reliability theory, the failure distribution represents a probabilistic description of the length of time during which a device is operative. One of the most frequently used distributions for this purpose is the exponential distribution. It has the memoryless property which characterises the failure distribution of devices which do not wear out with time. Hence it is crucial to reliability practitioners and researchers to test whether or not lifetimes of devices are exponentially distributed.

Against which class of failure distributions should we conduct a hypothesis testing of exponentiality? Some devices and structures deteriorate with time and it is useful to test if they show some ageing as opposed to random failures implying the exponential failure distribution.

The literature on testing for exponentiality is extensive. In particular, there are several different approaches to testing exponentiality against the increasing failure rate (IFR) alternative. For instance one can construct a test based on the total time on test transform or normalised spacings. For details the reader is referred to Barlow, et al. [2] and Doksum and Yandell [3].

In this paper, we mainly apply the results in Lo and Weng [7], which include those of Dykstra and Laud [4] as a special case.

The outline of this paper is: firstly our life testing model is described. Then we set a framework of a hypothesis testing of exponentiality vs IFRness from a nonparametric Bayesian perspective. In section 4, the weighted gamma process is introduced as our prior for the set of non-decreasing hazard rates. By an application of existing theorems, the posterior quantities of our interest are simplified. Section 5 includes a description of the weighted Chinese restaurant process, which is used to obtain a Monte carlo approximation to the posterior quantities of our interest.

2 Life testing models

Consider a life testing situation in which n new items are tested. Their lifetimes are assumed to be independent and identically distributed. Let $N \equiv \{N(t) : t \in [0, \infty)\}$ be a counting process representing the number of operative units at time t and the testing is terminated at a prespecified time $T \in (0, \infty)$. Let $x_1, \cdots, x_k \, (k \leq n)$ denote the failure times of n items on test, out of which k fail prior to or at T and the lifetimes of the remaining $n-k$ items are censored. Let $F(\cdot)$ be the cumulative distribution function for each item with density $f(\cdot)$. Then the survival function corresponding to $F(\cdot)$ and its hazard rate function are respectively denoted by

$$\begin{aligned} \bar{F}(t) &= 1 - F(t) \\ r(t) &= \frac{f(t)}{\bar{F}(t)} \end{aligned}$$

The likelihood function of $r(\cdot)$ given the data becomes

$$L\left(r(\cdot) \,|\, x_1, \cdots, x_k, Y(\cdot)\right) = \frac{n!}{(n-k)!} \left[\prod_{i=1}^{k} r(x_i)\right] \exp\left\{-\int I_{\{0 \leq s \leq T\}} Y(s) \times r(s) ds\right\}$$

where $I_{\{A\}}$ is the indicator function for set A and $Y(t)$ denotes the number of remaining items on test just prior to t (cf. Lo and Weng [7, Example 2.1 on page 230]).

We follow a model suggested by Lo and Weng [7], for which each hazard rate $r(\cdot)$ can be represented as a mixture of a known kernel with respect to some finite measure. Suppose that a nonnegative kernel $\kappa(t|\nu)$ on $([0,T] \times \Re, \mathcal{F} \times \mathcal{B})$ can be prespecified, where $\Re$ is the set of real numbers and $\mathcal{F}$ and $\mathcal{B}$ are Borel $\sigma-$fields of $[0,T]$ and $\Re$, respectively. Also assume that the following representation for $r(\cdot)$

$$r(t \,|\, \mu) = \int_{\Re} \kappa(t|\nu)\mu(d\nu) \;\; t \in [0,T]$$

where μ is a member of the space of finite measures on $(\Re, \mathcal{B})$ denoted by Θ and $\int r(t|\mu)dt$ is finite.

Under this mixture model for the hazard rates, the likelihood function is given by

$$L\left(\mu(\cdot) \,|\, x_1, \cdots, x_k, Y(\cdot)\right) = \frac{n!}{(n-k)!} \left[\prod_{1 \leq i \leq k} \int_{\Re} \kappa(x_i \,|\, \nu_i)\, \mu(d\nu_i)\right] \exp\left\{-\int_{\Re}\int_{\Re} I_{\{0 \leq s \leq T\}} Y(s)\, \kappa(s|\nu)\; ds \; \mu(d\nu)\right\}.$$

3 Hypothesis testing of exponentiality vs IFR-ness

Our main interest is to find evidence of the IFR property from failure data against their exponentiality. Null and alternative hypotheses are defined as follows.

$$\begin{array}{ll} H_0: & X_1, \cdots, X_n \text{ are } i.i.d. \text{ exponential.} \\ H_A: & X_1, \cdots, X_n \text{ are IFR, but not exponential.} \end{array}$$

Firstly, we assign prior probabilities to hypotheses denoted by $P(H_0)$ and $P(H_A)$. Secondly, Bayes theorem enables us to compute the posterior probability of each hypothesis. For instance, the posterior probability of the null hypothesis becomes

$$P(H_0 \mid x_1, \cdots, x_k, Y(\cdot)) = \frac{P(H_0) \int_0^\infty L(\lambda \mid x_1, \cdots, x_k, Y(\cdot))\pi(d\lambda)}{P(H_0) \int_0^\infty L(\lambda \mid x_1, \cdots, x_k, Y(\cdot))\pi(d\lambda) + P(H_A) \int_\Theta L(\mu(\cdot) \mid x_1, \cdots, x_k, Y(\cdot))G(d\mu(\cdot))}$$

where

$$L(\lambda \mid x_1, \cdots, x_k, Y(\cdot)) = \frac{n!}{(n-k)!}\lambda^k e^{-\lambda[\sum_{i=1}^k x_i + (n-k)T]}$$

and $\pi(d\lambda)$ and $G(d\mu(\cdot))$ denote priors for λ and $\mu(\cdot)$, respectively. Finally, one can make a decision by comparing $P(H_0 \mid x_1, \cdots, x_k, Y(\cdot))$ with $P(H_A \mid x_1, \cdots, x_k, Y(\cdot))$, namely, choose the hypothesis whose posterior probability is the largest. If a loss function is assessed, then the posterior expected loss with respect to each hypothesis needs to be evaluated.

As seen above, we need to compute

$$\int_0^\infty L(\lambda \mid x_1, \cdots, x_k, Y(\cdot))\pi(d\lambda) \tag{1}$$

$$\int_\Theta L(\mu(\cdot) \mid x_1, \cdots, x_k, Y(\cdot))G(d\mu(\cdot)). \tag{2}$$

In general, (1) can be easily evaluated. In particular, if $\pi(d\lambda)$ is a gamma distribution, then (1) is a Pareto density. On the other hand, evaluating (2) is quite difficult. We assume that a prior for $\mu(\cdot)$ is a gamma process and we will approximate (2) via a weighted Chinese restaurant process. Details are given in the following sections.

4 Prior distributions for hazard rates

Our choice of a prior for $\mu(\cdot)$ is a weighted gamma process, defined below.

Definition 4.1 Let $\alpha(s), s \geq 0$ be a function with $\alpha(0) \equiv 0$. A continuous time stochastic process $Z \equiv \{Z(s), s \geq 0\}$ is referred to as a gamma process with a shape function $\alpha(s), s \geq 0$ if the following properties are held.

1. $Z(0) \equiv 0$;

2. Z has independent increments;

3. for $t > s, Z(t) - Z(s) \sim gamma(\alpha(t) - \alpha(s), 1)$.

A generalisation of Z is defined next (cf. Dykstra and Laud [4, p. 357], Lo and Weng [7, pp. 231-232]). Let $\beta(s)$ be a nonnegative α-integrable function defined on $\Re$. A new continuous time stochastic process defined by

$$W(t) = \int_{[0,t)} \beta(s) dZ(s), \;\; t \in [0, \infty)$$

is referred to as an extended (weighted) gamma process with shape function $\alpha(\cdot)$ and scale function $\beta(\cdot)$.

We assume that $\{\mu(t), t \geq 0\}$ is an extended gamma process with shape function $\alpha(\cdot)$ and scale function $\beta(\cdot)$ with its distribution denoted by $G(d\mu(\cdot) \,|\, \alpha(\cdot), \beta(\cdot))$. Now we are ready to evaluate

$$\begin{aligned} & \int_\Theta L\left(\mu(\cdot) \,|\, x_1, \cdots, x_k, Y(\cdot)\right) G\left(d\mu(\cdot) | \alpha(\cdot), \beta(\cdot)\right) \\ & \quad = \int_\Theta \frac{n!}{(n-k)!} \left[\prod_{1 \leq i \leq k} \int_\Re \kappa(x_i \,|\, \nu_i)\, \mu(d\nu_i) \right] \cdot \\ & \qquad \exp\left\{ - \int_\Re \int_\Re I_{\{0 \leq s \leq T\}} Y(s)\, \kappa(s|\nu)\; ds\; \mu(d\nu) \right\} G\left(d\mu(\cdot) | \alpha(\cdot), \beta(\cdot)\right) \qquad (3) \end{aligned}$$

Several existing theorems can be applied to simplify (3). This is explained step by step below:

By Proposition 3.1 of Lo and Weng [7, pages 232–233], (3), which is also the Laplace transform of the prior for $\mu(\cdot)$, can be updated.

$$\begin{aligned} (3) \;\; = \;\; & \exp\left\{ - \int_\Re \log[1 + \beta(\nu) f(\nu)]\, \alpha(d\nu) \right\} \cdot \\ & \int_\Theta \frac{n!}{(n-k)!} \left[\prod_{1 \leq i \leq k} \int_\Re \kappa(x_i \,|\, \nu_i)\, \mu(d\nu_i) \right] G\left(d\mu(\cdot) \,\middle|\, \alpha(\cdot), \frac{\beta(\cdot)}{1 + \beta(\cdot) f(\cdot)} \right) \quad (4) \end{aligned}$$

where $f(\nu) = \int_\Re Y(s) \kappa(s|\nu) ds$. Then Lemma 3.1 of Lo and Weng [7, pages 233–234] is repeatedly applied to interchange the order of integrals as follows.

$$\begin{aligned} (4) \;\; = \;\; & C \int_\Re \int_\Theta \kappa(x_1|\nu_1) \left[\prod_{2 \leq i \leq k} \int_\Re \kappa(x_i \,|\, \nu_i)\, \mu(d\nu_i) \right] \\ & \qquad G\left(d\mu(\cdot) \,\middle|\, \alpha(\cdot) + \delta_{\nu_1}(\cdot), \frac{\beta(\cdot)}{1 + \beta(\cdot) f(\cdot)} \right) \cdot \frac{\beta(\nu_1)}{1 + \beta(\nu_1) f(\nu_1)}\, \alpha(d\nu_1) \\ = \;\; & C \int_\Re \int_\Re \int_\Theta \kappa(x_1|\nu_1) \kappa(x_2|\nu_2) \left[\prod_{3 \leq i \leq k} \int_\Re \kappa(x_i \,|\, \nu_i)\, \mu(d\nu_i) \right] \\ & \qquad G\left(d\mu(\cdot) \,\middle|\, \alpha(\cdot) + \delta_{\nu_1}(\cdot) + \delta_{\nu_2}(\cdot), \frac{\beta(\cdot)}{1 + \beta(\cdot) f(\cdot)} \right) \cdot \\ & \qquad \frac{\beta(\nu_2)}{1 + \beta(\nu_2) f(\nu_2)} (\alpha + \delta_{\nu_1})(d\nu_2) \cdot \frac{\beta(\nu_1)}{1 + \beta(\nu_1) f(\nu_1)}\, \alpha(d\nu_1) \\ & \vdots \end{aligned}$$

$$
\begin{aligned}
&\vdots \\
&= C \int_{\Re^k} \int_{\Theta} \left[\prod_{1\le i\le k} \kappa(x_i \mid \nu_i) \cdot \frac{\beta(\nu_i)}{1+\beta(\nu_i) f(\nu_i)} \right] \\
&\qquad G\left(d\mu(\cdot) \,\middle|\, \alpha(\cdot) + \sum_{1\le i\le k} \delta_{\nu_i}(\cdot), \frac{\beta(\cdot)}{1+\beta(\cdot)f(\cdot)} \right) \prod_{k\ge i\ge 1} \left(\alpha + \sum_{1\le j\le i-1} \delta_{\nu_j} \right)(d\nu_i) \\
&= C \int_{\Re^k} \left[\prod_{1\le i\le k} \kappa(x_i \mid \nu_i) \cdot \frac{\beta(\nu_i)}{1+\beta(\nu_i) f(\nu_i)} \right] \prod_{k\ge i\ge 1} \left(\alpha + \sum_{1\le j\le i-1} \delta_{\nu_j} \right)(d\nu_i) \qquad (5)
\end{aligned}
$$

where $(\alpha + \sum_{1\le j\le i-1} \delta_{\nu_j})(d\nu_i)$ denotes $\alpha(d\nu_1)$ when $i = 1$ and

$$
C = \frac{n!}{(n-k)!} \exp\left\{ - \int_{\Re} \log[1+\beta(\nu)f(\nu)]\, \alpha(d\nu) \right\}.
$$

(5) still has multiple integrals which make computation very difficult. However, through combinatorial techniques, (5) can be converted into a function of uni-dimensional integrals. Some notation is introduced before presenting a simplification of (5).

Consider a set of n elements, $\mathcal{S} = \{1, \ldots, n\}$. Let $\mathbf{p}$ and $n(\mathbf{p})$ be a partition of $\mathcal{S}$ and the number of cells of $\mathbf{p}$, respectively. C_i denotes the ith cell of $\mathbf{p}$ and e_i represents the number of elements in C_i. Applying Lemma 2 of Lo [6, page 353], we obtain

$$
(5) = C \sum_{\mathbf{p}} \prod_{i=1}^{n(\mathbf{p})} \left\{ (e_i - 1)! \int_{\Re} \prod_{l\in C_i} \left[\kappa(x_l \mid \nu) \cdot \frac{\beta(\nu)}{1+\beta(\nu)f(\nu)} \right] \alpha(d\nu) \right\} \qquad (6)
$$

where above terms are summed over all the possible partitions of $\{1, \ldots, k\}$. Since the case of non-decreasing failure rates corresponds to $\kappa(t|\nu) = I_{\{\nu\le t\}}$ (See Dykstra and Laud [4] and Lo and Weng [7, page 239]), we have

$$
\begin{aligned}
(6) &= C \sum_{\mathbf{p}} \prod_{i=1}^{n(\mathbf{p})} \left\{ (e_i - 1)! \int_{\Re} \prod_{l\in C_i} \left[I_{\{\nu\le x_l\}} \cdot \frac{\beta(\nu)}{1+\beta(\nu)f(\nu)} \right] \alpha(d\nu) \right\} \\
&= C \sum_{\mathbf{p}} \prod_{i=1}^{n(\mathbf{p})} \left\{ (e_i - 1)! \int_{\Re} \left[\frac{\beta(\nu)}{1+\beta(\nu)f(\nu)} \right]^{e_i} I_{\{0\le\nu\le\min(i)\}} \alpha(d\nu) \right\} \qquad (7)
\end{aligned}
$$

where $\min(i) = \min\{x_l : l \in C_i\}$.

When $|\mathcal{S}|$ is large, a direct computation of (7) is intractable, hence an approximation method is required to evaluate it. The next section describes a weighted Chinese restaurant process, which enables one to approximate (7).

5 Monte Carlo approximations to the posterior probabilites

The Chinese restaurant process is a random mechanism to sequentially partition a group of “people” denoted by $\mathcal{S} = \{1, \ldots, n\}$ into “tables” denoted by $\{C_1, \ldots, C_{n(\mathbf{p})}\}$

where $n(\mathbf{p})$ denotes the number of tables, i.e., subgroups (see Aldous [1], Kuo [5]). An importance sampling variant of it called the weighted Chinese restaurant process is used to approximate (7), which takes into account of similarities in failure times, x_js. The following description of the Chinese restaurant process is based on Lo, Brunner and Chan [8] to which the reader is referred for details.

Let $\rho(C)$ define the marginal weight for a table C by

$$\rho(C) = \int \prod_{j \in C} w_j(\nu)\alpha(d\nu)$$

where $w_j(u)$ is a nonnegative finite "likelihood" weighting function for $j \in \mathcal{S}$ and $\alpha(du)$ is a "prior" mixing major. We also define a "predictive" weights of $r \notin C$ given C by the ratio

$$\rho(r|C) = \begin{cases} \rho(r, C)/\rho(C) & \text{if } \rho(C) \neq 0 \\ 0 & \text{if } \rho(C) = 0. \end{cases}$$

The weighted Chinese restaurant process proceeds as follows.

Step 1. Set $\lambda(0) = \rho(1)$. Assign 1 to C_1 with probability $\rho(1)/\lambda(0) = 1$.

Step r $(r = 2, \ldots, n)$. From Step $r-1$, we have tables $C_1, \ldots, C_{n(\mathbf{p})}$ with their respective sizes $e_1, \ldots, e_{n(\mathbf{p})}$.

- Compute $\lambda(r-1) = \rho(r) + \sum_{1 \leq i \leq n(\mathbf{p})} e_i \rho(r|C_i)$.
- Assign r to a new table $C_{n(\mathbf{p})+1}$ with probability $\rho(r)/\lambda(r-1)$; otherwise, assign r to C_i with probability $e_i\rho(r|C_i)/\lambda(r-1), i = 1, \ldots, n(\mathbf{p})$.
- If r is assigned to a new table, $n(\mathbf{p}) \leftarrow n(\mathbf{p}) + 1$; otherwise $n(\mathbf{p})$ stays the same.

A density of the above algorithm is given by

$$q(\mathbf{p}|\alpha, \mathbf{w}) = \phi(\mathbf{p})/\Lambda_{n-1}$$

where

$$\begin{aligned} \phi(\mathbf{p}) &= \prod_{1 \leq i \leq n(\mathbf{p})} (e_i - 1)!\rho(C_i) \\ \Lambda_{n-1}(\mathbf{p}) &= \prod_{j=0}^{n-1} \lambda(j) \\ \mathbf{w} &= (w_1(\cdot), \ldots, w_n(\cdot)). \end{aligned}$$

If we let $w_j(\nu) = I_{\{0 \leq \nu \leq x_j\}}$, then

$$\begin{aligned} \phi(\mathbf{p}) &= \prod_{1 \leq i \leq n(\mathbf{p})} (e_i - 1)!\rho(C_i) \\ &= \prod_{1 \leq i \leq n(\mathbf{p})} (e_i - 1)! \int \prod_{j \in C_i} w_j(\nu)\alpha(d\nu) \end{aligned}$$

$$= \prod_{1\le i\le n(\mathbf{p})} (e_i - 1)! \int \prod_{j\in C_i} I_{\{0\le\nu\le x_j\}}\alpha(d\nu)$$
$$= \prod_{1\le i\le n(\mathbf{p})} (e_i - 1)! \int I_{\{0\le\nu\le \min(i)\}}\alpha(d\nu).$$

Rewriting (7) we get

$$C\sum_{\mathbf{p}}\left\{\prod_{i=1}^{n(\mathbf{p})}\left[(e_i-1)!\int_{\Re} I_{\{0\le\nu\le\min(i)\}}\alpha(d\nu)\right]\times\prod_{i=1}^{n(\mathbf{p})}\int_{\Re}\left[\frac{\beta(\nu)}{1+\beta(\nu)f(\nu)}\right]^{e_i}\alpha(d\nu\,|\,C_i)\right\} \tag{8}$$

where

$$\alpha(d\nu\,|\,C_i) = \frac{I_{\{0\le\nu\le\min(i)\}}\alpha(d\nu)}{\int_{\Re} I_{\{0\le\nu\le\min(i)\}}\alpha(d\nu)}.$$

Generate sufficiently large number of partitions $\mathbf{p}_1,\cdots,\mathbf{p}_M$ according to the density $q(\mathbf{p}|\alpha,\mathbf{w})$ of the weighted Chinese restaurant process. Then an approximation of (8) is given by

$$\frac{1}{M}\sum_{k=1}^{M} h(\mathbf{p}_k)\Lambda_{n-1}(\mathbf{p}_k)$$

where

$$h(\mathbf{p}) = \prod_{i=1}^{n(\mathbf{p})}\int_{\Re}\left[\frac{\beta(\nu)}{1+\beta(\nu)f(\nu)}\right]^{e_i}\alpha(d\nu\,|\,C_i).$$

6 Concluding remarks

This paper presents a test to detect IFR failure data against no-ageing data, i.e. exponentiality of failure data. The gamma process is used as a prior for the set of non-decreasing failure functions. The posterior quantities are obtained by approximating integrals via weighted Chinese restaurant processes. The approach taken here is straightforward and a decision is based on the posterior probability of hypothesis. This is in contrast to some methods suggested in the literature based on the (asymptotic) distribution of test statistics, methods which violate the Likelihood Principle.

Acknowledgments

Part of this study was done while the first author was visiting Albert Y. Lo at the Department of Information and Systems Management, Hong Kong University of Science & Technology. The authors are extremely grateful to him for helpful discussions and his assistance. Also many thanks to Ray Brownrigg for his help.

References

[1] Aldous, D. J. Exchangeability and Related Topics. In *École d'Été de Probabilités de Saint-Flour XIII - 1983*, pages 1–198. Springer-Verlag, 1985.

[2] Barlow, R.E., Bartholomew, D.J., Bremner, J.M. and Brunk, H.D. *Statistical Inference under Order Restrictions: The Theory and Applications of Isotonic Regression.* John Wiley & Sons, 1980.

[3] Doksum, K.A. and Yandell, B.S. Tests for Exponentiality. In *Handbook of Statistics*, volume 4, pages 579–611. Elsevier Science Publishers, 1984.

[4] Dykstra, R.L. and Laud, P. A Bayesian nonparametric approach to reliability. *The Annals of Statistics*, 9:356–367, 1981.

[5] Kuo, L. Computaions of mixtures of Dirichlet processes. *SIAM J. Sci. Statist. Comput.*, 7:60–71, 1986.

[6] Lo, A.Y. On a class of Bayesian nonparametric estimates: I. density estimates. *The Annals of Statistics*, 12(1):351–357, 1984.

[7] Lo, A.Y. and Weng, C. On a class of Bayesian nonparametric estimates: II. hazard rate estimates. *Ann. Inst. Statist. Math.*, 41(2):227–245, 1989.

[8] Lo, A.Y., Brunner, L.J. and Chan, A.T. Weighted Chinese restaurant processes and Bayesian mixture models. research manuscript, 1999.