CONDITIONAL LAPLACE TRANSFORMS FOR BAYESIAN NONPARAMETRIC INFERENCE IN RELIABILITY THEORY

Larry P. AMMANN
University of Texas at Dallas, P.O. Box 830688, Richardson, TX 75083-0688, USA

Received 21 November 1983
Revised 12 March 1985

In order to apply Bayesian nonparametric methods to reliability problems, it is desirable to have available priors over a broad class of survival distributions. In the present paper, this is achieved by taking the failure rate function to be the sum of a nonnegative stochastic process with increasing sample paths and a process with decreasing sample paths. This approach produces a prior which chooses an absolutely continuous survival distribution that can have an IFR, DFR, or U-shaped failure rate. Posterior Laplace transforms of the failure rate are obtained based on survival data that allows censoring. Bayes estimates of the failure rate as well as the lifetime distribution are then calculated from these posterior Laplace transforms. This approach is also applied to a competing risks model and the proportional hazards model of Cox.


Bayesian nonparametric estimation * conditional Laplace transform * competing risks * proportional hazards

1. Introduction

Bayesian nonparametric inference has received much attention in the recent literature. With this approach, the statistician selects a prior over a space of suitable probability measures. Ferguson (1973) and others have used the Dirichlet process as a prior since it is a stochastic process with sample realizations which are probability measures, and the posterior process is also a Dirichlet process. Since these realizations are almost surely purely atomic, the Dirichlet process provides a prior over a space of purely atomic probability measures. Doksum (1974) and Ferguson and Phadia (1979) constructed another class of priors in which the hazard function,

\[ H(t) = -\log[1 - F(t)], \]

of a lifetime distribution \( F \) is a nonnegative stochastic process with independent increments. Such processes are also purely atomic, and so this construction again produces a class of priors over a space of purely atomic probability measures.

In order to construct priors over a space of absolutely continuous (with respect to Lebesgue measure) probability measures for some reliability applications, Dykstra...
and Laud (1981) and Lo (1984) represent the failure rate,

\[ h(t) = \frac{\partial}{\partial t} H(t), \]

as a gamma process. Since the gamma process has nondecreasing sample paths, this approach results in a prior over a space of IFR survival distributions. The aim of the present paper is to put the approach of Dykstra and Laud and Lo in a more general setting by representing the hazard rate as a function of the sample paths of nonnegative processes with independent increments which consists of an increasing component and a decreasing component. Ammann (1984) uses this approach to construct a model for Bayesian nonparametric estimation of a tolerance distribution. This results in a broad class of priors over a space of absolutely continuous probability measures which contains IFR, DFR, and U-shaped failure rate survival distributions. Posterior Laplace transforms of these processes are obtained for several reliability applications. These Laplace transforms enable one to find Bayes estimates under various loss functions directly, without the need to resort to the limiting arguments of Dykstra and Laud and Lo.

In Section 2, the priors are defined for the model described above, and posterior Laplace transforms are obtained for data that may contain censored observations. These Laplace transforms are then used to find Bayes estimates for two different loss functions.

The results of Section 2 are used in Section 3 to develop a Bayesian nonparametric approach for competing risks. Competing risks models have been applied in a wide variety of areas. For example, these models are useful in the study of the reliability of a system of components each of which is subject to failure, and of life-lengths of biological systems that are susceptible to several causes of death. In this section, posterior Laplace transforms are given for competing risks data, and some estimation problems are discussed.

The results of Section 2 are also extended in Section 4 to the proportional hazards model of Cox (1972). Wild and Kalbfleisch (1981) discuss an extension of Ferguson and Phadia for this situation. Here the baseline failure rate is taken to be a nonnegative process with independent increments. A likelihood function is given based on the marginal survival probability, which can then be used to estimate covariate parameters. The posterior Laplace transform of the failure rate is obtained when the covariate parameters are known (or estimated). Bayes (or partial Bayes) estimates are then computed from this Laplace transform.

2. Survival distributions with completely random failure rates

In this section, priors are constructed over a space of absolutely continuous probability measures on \( T = [0, \infty) \) by taking the failure rate function to be the sum of two independent, nonnegative stochastic processes with independent increments.
Such processes generate random measures, called completely random measures by Kingman (1967). No notational distinction will be made here between the process with independent increments and the generated completely random measure.

It is well-known that completely random measures are infinitely divisible, and so their Laplace transforms have a special form given by

$$\Psi(\xi) = E \exp\left\{ -\int_T \xi(t) Y(dt) \right\}$$

$$= \exp\left\{ -\int_T \xi(t) \lambda(dt) + \int_T \int_{\mathbb{R}_+} (e^{-\nu(t)} - 1) Q(du, dt) \right\},$$

where $\lambda$ is a locally finite measure, $T = [0, \infty)$, $\mathbb{R}_+ = (0, \infty)$, $a \wedge b = \min(a, b)$ and $Q$ is a measure on $T \times \mathbb{R}_+$ that satisfies

$$\int_{\mathbb{R}_+} (\nu \wedge 1) Q(du, A) < \infty \quad \text{for every compact measurable set } A \subset T. \quad (2.1)$$

The Laplace transform is defined on the collection of all bounded, measurable functions $\xi$ that vanish outside compact sets. It is assumed throughout that all completely random measures considered are centered, i.e., $\lambda = 0$. The measure $Q$ is called the Lévy measure of $Y$. It governs the locations and masses of the atoms of $Y$. In particular, $Q(B, A)$ represents the expected number of atoms of $Y$ that are located within $A$ and have masses in $B$.

Condition (2.1) implies that the total mass of all atoms in a compact set $A$ is finite a.s., although the total number of all atoms in $A$ may be countably infinite. The gamma process is an example of a completely random measure in which the marginal distribution of $Y(A)$ is a gamma distribution. The Lévy measure of the gamma process is given by

$$Q(du, dt) = \nu^{-1} e^{-\nu/\beta(t)} du \alpha(dt),$$

where $\beta(t)$ is a strictly positive, measurable function called the scale function, and $\alpha$ is a locally finite measure called the shape measure. This process places a finite or countably infinite number of atoms within any compact set $A$ for which $\alpha(A) > 0$.

Let $B$ denote the class of all completely random measures on $T$ and let $B^*$ denote the subclass of $B$ which contains all completely random measures whose Lévy measures satisfy

$$\int_{\mathbb{R}_+} (\nu \wedge 1) Q(du, T) < \infty. \quad (2.2)$$

Note that (2.2) implies that $Y(T) < \infty$ a.s., i.e., $Y(T)$ is a real-valued random variable.

Next let $Y_1 \in B$, $Y_2 \in B^*$ with Lévy measures $Q_1$, $Q_2$ respectively, and define

$$h(t) = Y_1[0, t] + Y_2(t, \infty), \quad F_h(t) = 1 - \exp\left\{ -\int_0^t h(s) \, ds \right\}.$$
It is easily seen that if $Q_1$ is not the zero measure or if

$$\int_T \int_{\mathbb{R}_+} (1 - e^{-\nu}) Q_2(du, dt) = \infty,$$  \hfill (2.3)

then $F_h$ is a proper distribution function a.s. Let $X$ denote a random variable whose conditional distribution function given $h$ is $F_h$. Then $X$ can be thought of as the lifetime of a device with unknown failure rate $h$. A prior for $h$ is specified by selection of appropriate Lévy measures $Q_1$ and $Q_2$. This construction can be interpreted in the following way. The component of $h$ given by $Y_1$ represents environmental damage the device will experience during its lifetime. This component causes an increase in the failure rate as time progresses. The component of $h$ given by $Y_2$ represents internal and/or external hazards which the device can learn to overcome. The rate at which these hazards can be removed is specified by $Q_2$. Note that if $Q_2$ is the zero measure, then this prior chooses an IFR distribution for $X$, and if $Q_1$ is the zero measure, then this prior chooses a DFR distribution for $X$. Also, if

$$Q_1(du, [0, t_1]) = 0 \quad \text{and} \quad Q_2(du, [t_2, \infty)) = 0,$$

for some $t_1 \leq t_2$, then $h$ is decreasing on $[0, t_1]$ and increasing on $[t_2, \infty)$.

Now suppose that $X_1, \ldots, X_n$ represents a set of random variables that are conditionally i.i.d. given $h$ with conditional distribution function $F_h$. Also, let $C_1, \ldots, C_n$ denote a set of i.i.d. censoring variables that are independent of $X_1, \ldots, X_n, h$ and have absolutely continuous distribution function $G$ with density $g$. The goal here is to estimate $h$ based on $Z_1, \ldots, Z_n, A_1, \ldots, A_n, \Delta_1, \ldots, \Delta_n$, where

$$Z_i = \min(X_i, C_i), \quad \Delta_j = I\{X_j \leq C_j\}.$$

To this end, let $\mu(\cdot; \xi_1, \xi_2)$ denote a measure on $T^n \times \{0, 1\}^n$ generated by

$$EP(Z_i > u_i, \Delta_1 = \delta_1, \ldots, Z_n > u_n, \Delta_n = \delta_n | Y_1, Y_2) \cdot \exp\left\{-\int_T \xi_1(t) Y_1(dt) - \int_T \xi_2(t) Y_2(dt)\right\}.$$

Note that $\mu$ characterizes the joint probability measure of

$$Z_1, \Delta_1, \ldots, Z_n, \Delta_n, Y_1, Y_2,$$

and that the joint posterior Laplace transform of $Y_1, Y_2$ is the Radon–Nikodym derivative,

$$\frac{d\mu(\cdot; \xi_1, \xi_2)}{d\mu(\cdot; 0, 0)}.
$$  \hfill (2.4)

Set $m = \sum_{j=1}^n \delta_j$. Let $s_1, \ldots, s_m$ denote the uncensored observations and let $w_1, \ldots, w_{n-m}$ denote the censored observations. It is assumed throughout that each
of these sets of measurements is ordered. Then

\[ \mu(A; \xi_1, \xi_2) = \int_A \left[ \prod_{i=1}^m \mathcal{G}(s_i) \right] \left[ \prod_{j=1}^{n-m} g(w_j) \right] \cdot E \left[ \prod_{i=1}^m h(s_i) \exp \left\{ - \int_{0}^{s_i} h(t) \, dt \right\} \prod_{i=1}^m ds_i \right. \]

\[ \cdot \exp \left\{ - \sum_{j=1}^{n-m} \int_{0}^{w_j} h(t) \, dt - \int_{\tau}^{\xi_1(t)} Y_1(dt) - \int_{\tau}^{\xi_2(t)} Y_2(dt) \right\} \]. \quad (2.5) \]

Hence, from (2.4) and (2.5), the joint posterior Laplace transform of \( Y_1 \) and \( Y_2 \) is given by

\[ E \prod_{i=1}^m h(s_i) \exp \left\{ - \int_{0}^{s_i} h(t) \, dt \right\} \]

\[ \cdot \exp \left\{ - \sum_{j=1}^{n-m} \int_{0}^{w_j} h(t) \, dt - \int_{\tau}^{\xi_1(t)} Y_1(dt) - \int_{\tau}^{\xi_2(t)} Y_2(dt) \right\} \]

\[ \cdot \left( E \prod_{i=1}^m h(s_i) \exp \left\{ - \int_{0}^{s_i} h(t) \, dt \right\} \exp \left\{ - \sum_{j=1}^{n-m} \int_{0}^{w_j} h(t) \, dt \right\} \right)^{-1}. \quad (2.6) \]

As should be expected, this posterior Laplace transform does not depend on the censoring distribution \( G \). To evaluate (2.6), first note that by Tonelli's Theorem (Royden (1968, p. 270)),

\[ \int_{0}^{u} h(t) \, dt = \int_{\tau}^{(u-t)^+} Y_1(dt) + \int_{\tau}^{(u \land t)} Y_2(dt), \quad (2.7) \]

where \( a^+ = \min(a, 0) \). The posterior failure rate is first given for the case in which all observations are censored. It is obtained directly from (2.6) and (2.7).

**Theorem 2.1.** If \( m = 0 \), then the posterior failure rate \( h \) is given by

\[ h(t; w) = Y_{1,w}[0, t] + Y_{2,w}(t, \infty), \]

where \( Y_{1,w} \) and \( Y_{2,w} \) are independent, completely random measures with Lévy measures, respectively,

\[ Q_1(du, dt; w) = \exp \left\{ -u \sum_{j=1}^{n} (w_j - t)^+ \right\} Q_1(du, dt), \]

\[ Q_2(du, dt; w) = \exp \left\{ -u \sum_{j=1}^{n} (w_j \land t) \right\} Q_2(du, dt). \]
If $Y_i$ is a gamma process with shape $\alpha$ and scale function $\beta(t)$, then $Y_{1,\omega}$ is a gamma process with shape $\alpha$ and scale function given by

$$\beta_w(t) = \beta(t) \left[ 1 + \beta(t) \sum_{j=1}^{n} (w_j - t)^+ \right]^{-1}. $$

This result for gamma processes is Theorem 3.2 of Dykstra and Laud (1981), but is obtained here directly via the Laplace transform without the need for their limiting arguments.

Next set

$$f_1(t; w, s) = \sum_{i=1}^{m} (s_i - t)^+ + \sum_{j=1}^{n-m} (w_j - t)^+$$

and

$$f_2(t; w, s) = \sum_{i=1}^{m} (s_i \wedge t) + \sum_{j=1}^{n-m} (w_j \wedge t).$$

Then, for $m > 0$,

$$E \prod_{i=1}^{m} h(s_i) \exp \left\{ - \int_{0}^{h} h(t) \, dt \right\} \cdot \exp \left\{ - \sum_{i=1}^{m} \int_{0}^{h} \xi_1(t) \, dt - \int_{T} \xi_2(t) \, Y_1(dt) \right\}$$

$$= \frac{\partial^m}{\partial \theta_1 \cdots \partial \theta_m} \left[ E \exp \left\{ - \int_{T} \left[ \sum_{i=1}^{m} \Theta_i I(t \leq s_i) + f_1(t; w, s) + \xi_1(t) \right] Y_1(dt) \right\} \right]_{\theta = 0}$$

$$= \frac{\partial^m}{\partial \Theta_1 \cdots \partial \Theta_m} \left[ \exp \left\{ \int \left( \exp \left\{ -v \left[ \sum_{i=1}^{m} \Theta_i I(t \leq s_i) + f_1(t; w, s) + \xi_1(t) \right] \right\} - 1 \right) Q_1(du, dt) \right. \right.$$

$$+ \left. \int \left( \exp \left\{ -v \left[ \sum_{i=1}^{m} \Theta_i I(t > s_i) + f_2(t; w, s) + \xi_2(t) \right] \right\} - 1 \right) Q_2(du, dt) \right\} \right|_{\theta = 0}.$$

(2.8)

Now let $\Gamma = \Gamma(m)$ denote the set of all distinct partitions of the integers $1, \ldots, m$, let $\sigma$ denote an element of $\Gamma$, and let $\tau$ denote a group in $\sigma$. With this notation, an expression for (2.6) can be obtained when $m > 0$ by application of (2.8) to (2.6) after straightforward, although tedious, calculations to evaluate (2.8).

**Theorem 2.2.** If $m > 0$, then the joint posterior Laplace transform of $Y_1$ and $Y_2$ is
given by
\[
\Psi(\xi_1, \xi_2; w, s) = q(w, s; \xi_1, \xi_2)[q(w, s; 0, 0)]^{-1}
\cdot \exp \left\{ \iint (e^{-\xi_1(t)} - 1) \exp\{-\psi_1(t; w, s)\} Q_1(dv, dt) \right. \\
+ \left. \iint (e^{-\xi_2(t)} - 1) \exp\{-\psi_2(t; w, s)\} Q_2(dv, dt) \right\},
\]
where
\[
q(w, s; \xi_1, \xi_2) = \sum_{\sigma \in I} \prod_{\tau \in \sigma} \left\{ \iint I\{t \leq \min_{i \in \tau} s_i\} v^j e^{-\psi_1(t; w, s) + \psi_1(t)} Q_1(dv, dt) \\
+ \iint I\{t > \max_{i \in \tau} s_i\} v^j e^{-\psi_1(t; w, s) + \psi_1(t)} Q_2(dv, dt) \right\}.
\] (2.9)

To complete this section, Theorem 2.2 is applied to an estimation problem. Suppose that a statistician’s loss function reflects a desire to estimate the failure rate \(h\), e.g.,
\[
L(h, \hat{h}) = \int [h(t) - \hat{h}(t)]^2 W(dt),
\]
where \(W\) is a measure that satisfies
\[
\int \text{var} Y_1[0, t] W(dt) < \infty \quad \text{and} \quad \int \text{var} Y_2(t, \infty) W(dt) < \infty.
\]
In this case, the Bayes estimate of \(h\) is the posterior mean of \(h\). Set \(\xi_1(t) = \Theta I\{t \leq u\}\) and \(\xi_2(t) = \Theta I\{t > u\}\). Then
\[
\hat{h}(u) = -\frac{\partial}{\partial \Theta} \Psi(\xi_1, \xi_2; w, s)|_{\Theta = 0}.
\]

**Theorem 2.3.** If the loss function satisfies (2.10), and if \(m = 0\), then
\[
\hat{h}(u) = \int_0^u \int_{R^+} v Q_1(dv, dt; w) + \int_u^\infty \int_{R^+} v Q_2(dv, dt; w).
\]
If \(m > 0\), then
\[
\hat{h}(u) - q(w, s(u); 0, 0)/q(w, s; 0, 0),
\]
where \(s(u)\) denotes the ordered values of \(u, s_1, \ldots, s_m\).

**Example.** Let \(Y_1\) and \(Y_2\) be independent gamma processes with Lévy measures
\[
Q_1(dv, dt) = \alpha(v^{-1} e^{-v/\beta} dv \ dt),
\]
\[
Q_2(dv, dt) = \begin{cases} 
\lambda v^{-1} e^{-v/\delta} dv \ dt, & 0 \leq t \leq 1, \\
0, & t > 1,
\end{cases}
\]
where $\alpha, \beta, \lambda, \delta > 0$. In this case, the prior expected failure rate, $h_0(t)$, is given by

$$h_0(t) = EY_1[0, t] + EY_2(t, \infty) = \begin{cases} \alpha \beta t + \lambda \delta (1-t), & 0 \leq t \leq 1, \\ \alpha \beta t, & t > 1. \end{cases}$$

Thus, if $\lambda \delta > \alpha \beta$, then $h_0(t)$ initially decreases but eventually increases. In order to obtain the Bayes estimate for this example, it is necessary to evaluate the integrals:

$$\int_0^u \int_{R_+} v' e^{-v_1(t; w, \omega)} Q_1(dt, dv), \quad (2.11)$$

$$\int_u^1 \int_{R_+} v' e^{-v_2(t; w, \omega)} Q_2(dt, dv). \quad (2.12)$$

To evaluate (2.11), first let $u$ denote an arbitrary positive real number and let $u_1, \ldots, u_n$ denote the ordered values of $s_1, \ldots, s_m, w_1, \ldots, w_{n-m}$. If $u_{k-1} < u \leq u_k$, then

$$\int_0^u \int_{R_+} v' e^{-v_1(t; w, \omega)} Q_1(dt, dv) = \int_0^{u_{k-1}} \left[ 1 + \beta \sum_{i=r}^n (u_i - t) \right]^{-j} dt + \int_{u_{k-1}}^u \left[ 1 + \beta \sum_{i=r}^n (u_i - t) \right]^{-j} dt,$$

and if $u > u_n$, then

$$\int_0^u \int_{R_+} v' e^{-v_1(t; w, \omega)} Q_1(dt, dv) = \alpha (j-1)! \beta^j \left\{ \sum_{i=r}^n \int_{u_{i-1}}^{u_i} \left[ 1 + \beta \sum_{i=r}^n (u_i - t) \right]^{-j} dt \right\}.$$

To evaluate (2.12), let $0 < u < 1$, and suppose that $u_{a-1} < u \leq u_a$ and $u_{k-1} \leq u < u_k$ for some $1 \leq k \leq a \leq n$. Then

$$\int_0^u \int_{R_+} v' e^{-v_2(t; w, \omega)} Q_2(dt, dv) = \lambda (j-1)! \delta^j \left\{ \int_u^{u_a} \left[ 1 + \delta \sum_{i=1}^{k-1} u_i + \delta (n-k+1) t \right]^{-j} dt \right\} + \sum_{r=k}^{a-1} \int_{u_r}^{u_{r+1}} \left[ 1 + \delta \sum_{i=1}^r u_i + \delta (n-r) t \right]^{-j} dt.$$

These expressions can then be substituted into the results of Theorem 2.3 to obtain the Bayes estimate of the failure rate.

The results given here are based on simple life-testing data in which each device is tested separately. However, they can also be applied to renewal life-testing, which would give a combination of complete and censored data. Furthermore, these results can also be applied to a Bayesian nonparametric approach for age-dependent branching processes.

### 3. Competing risks

This section discusses a general framework for Bayesian nonparametric inference for a competing risks model when the prior chooses the failure rates of the survival
distribution. Suppose that a system is composed of \( k \) components in series so that the system fails when one of the components fails. Let \( Z_1, \ldots, Z_k \) denote the component failure times and let \( X \) denote the system failure time. Then

\[
X = \min(Z_1, \ldots, Z_k).
\]

Also, let \( \delta \) denote the component that causes system failure. That is,

\[
\{X \leq t, \delta = j\} = \{Z_j \leq t, Z_i < Z_j, i \neq j\}.
\]

The basic goal in this situation is to make inferences regarding the marginal behavior of \( \{Z_i\} \) based on \( (X, \delta) \).

Let \( h_j(t) \) denote the marginal failure rate function of \( Z_j, 1 \leq j \leq k \), and let \( J \) denote a subsystem of components. It is assumed that the subsystem aggregate failure rate function \( h_J(t) \) is given by

\[
h_J(t) = \sum_{j \in J} h_j(t).
\]

This assumption implies that the component failure times are independent given the failure rates \( \{h_j(t)\} \). In this case the probability that the system survives past time \( t \) but eventually fails due to component \( j \) is given by

\[
P(X > t, \delta = j) = P(t < Z_j < Z_i, i \neq j)
\]

\[
= \int_t^\infty h_j(s) \exp\left\{ -\sum_{i=1}^k \int_0^s h_i(u) \, du \right\} \, ds.
\]

As noted by Prentice et al (1978), the individual failure rates \( \{h_j(t)\} \) are identifiable from the marginal distribution of \( (X, \delta) \).

This model can be put into a nonparametric Bayesian context by allowing the failure rates to be functions of stochastic processes for which, with probability one, the stochastic integrals,

\[
\int_T \xi(t) h_j(dt), \quad 1 \leq j \leq k,
\]

exist for bounded measurable functions \( \xi \) that vanish outside compact sets. The joint Laplace transform of such stochastic processes is then defined to be

\[
\Psi(\xi_1, \ldots, \xi_k) = E \exp\left\{ -\sum_{j=1}^k \int_T \xi_j(t) h_j(dr) \right\}.
\]

This Laplace transform characterizes the joint probability measure of the failure rates.

Now suppose that \( (X_1, \delta_1), \ldots, (X_n, \delta_n) \) are i.i.d. observations given failure rates. Then

\[
P(X_r > t_r, \delta_r = i_r, 1 \leq r \leq n | h_j, 1 \leq j \leq k)
\]

\[
= \prod_{r=1}^n \int_{t_r}^\infty \prod_{i=1}^n h_i(s_r) \exp\left\{ -\sum_{j=1}^k \int_0^{s_r} h_j(u) \, du \right\} \, ds_r.
\]
As in Section 2, define a measure $\mu$ on $[0, \infty)^n$ by

$$
\mu(A; \xi_1, \ldots, \xi_k) = E \left[ \prod_{r=1}^{n} \int_A \left[ \sum_{j=1}^{k} h_j(s_r) \exp\left\{ - \sum_{j=1}^{k} \int_0^{s_r} h_j(u) \, du \right\} \, ds_r \right] \right],
$$

$$
\exp\left\{ - \sum_{j=1}^{k} \int_T \xi_j(u) h_j(du) \right\}. \quad (3.2)
$$

Note that $\mu$ characterizes the joint probability measure of

$$(X_1, \delta_1), \ldots, (X_m, \delta_n), \quad h_1, \ldots, h_k,$$

and the posterior joint Laplace transform of $h_1, \ldots, h_k$ given

$$(X_1, \delta_1), \ldots, (X_m, \delta_n)$$

is the Radon–Nikodym derivative

$$
\Psi(\xi_1, \ldots, \xi_k | (X_1, \delta_1), \ldots, (X_m, \delta_n)) = \frac{d\mu(\cdot; \xi_1, \ldots, \xi_k)}{d\mu(\cdot; 0, \ldots, 0)}. \quad (3.3)
$$

The techniques developed in Section 2 can be utilized to construct a competing risks model with random failure rates that can be increasing, decreasing, or U-shaped. In order to illustrate this application, the priors examined in this section will choose increasing failure rates. The extension to the general case should be apparent.

Let $Y_i, 1 \leq i \leq k$, be independent completely random measures with Lévy measures $Q_i$ respectively, where $Q_i$ satisfies (2.2), $1 \leq i \leq k$. Next define the failure rates $h_i(t)$ by

$$
h_i(t) = Y_i[0, t]. \quad (3.4)
$$

Next suppose that $n$ identical, independent systems are tested and that the failure time and cause of failure are observed for each system. Then the observations have the form

$$(X_1 = x_1, \delta_1), \ldots, (X_n = x_n, \delta_n).$$

Define $n_i = \sum_{r=1}^{n} I\{\delta_r = i\}$ and let $D = \{i: n_i > 0\}$. For each $i \in D$, let $X_{ir}, 1 \leq r \leq n_i$, denote the ordered failure times of those systems that fail due to component $i$, and let $\gamma_{ir}(s) = 1$ if $s \leq x_{ir}$ and $\gamma_{ir} = 0$ otherwise. In order to obtain the posterior Laplace transform of the failure rates given by (1.5), it is necessary to evaluate (3.3).

This posterior Laplace transform can be obtained using similar techniques to those used in Section 2. Some additional notation is required to obtain (3.4). For each $1 \leq i \leq k$, define $I_i = I_i(n_i)$ to be the set of all distinct partitions of the integers $1, \ldots, n_i$, let $\sigma_i$ denote an element of $I_i$, and let $\tau_i$ denote a group in $\sigma_i$. The posterior Laplace transform of the failure rates is given in Theorem 3.1. Its proof is similar to the proof of Theorem 2.2, and so is omitted.
Theorem 3.1. The posterior Laplace transform of

\( h_1, \ldots, h_k \)

is given by

\[
\Psi(\xi_1, \ldots, \xi_k | X_1 = x_1, \delta_1, \ldots, X_n = x_n, \delta_n) = \prod_{i \in D} [q_i(x; \xi_1, \ldots, \xi_k)] [q_i(x; 0, \ldots, 0)]^{-1} 
\cdot \exp \left\{ \sum_{j=1}^k \int T \int_{\mathbb{R}_+} (e^{-\varepsilon_j(s)} - 1) Q_j(dv, ds; x) \right\},
\]

where

\[
q_i(x; \xi_1, \ldots, \xi_k) = \sum_{\sigma_j \in F_j, \tau \in \sigma} \int T \int_{\mathbb{R}_+} I \left\{ s \leq \min_{r \in \tau}(x_{ir}) \right\} v^j e^{-\varepsilon_j(s)} Q_i(dv, ds),
\]

\[
\varepsilon_j^g(s) = \sum_{i \in D} \sum_{r=1}^n (x_{ir} - s)^+ + \xi_j(s),
\]

\[
Q_j(dv, ds; x) = \exp \left\{ -v \sum_{i \in D} \sum_{r=1}^n (x_{ir} - s)^+ \right\} Q_j(dv, ds).
\]

Next suppose that some observations are censored. In practice, censored observations might arise from the termination of testing or from a system that fails due to some cause other than the \( k \) causes under study. If the censoring is due to some random mechanism, then the censoring variable can be treated as an additional cause of failure and studied along with the other causes by application of Theorem 3.1. What will be considered here is a deterministic censoring mechanism. Let \( Z_1, \ldots, Z_m \) denote the (unobserved) failure times of the systems that are censored, and let \( z_1, \ldots, z_m \) denote the censoring times of these systems. Then this data has the form

\( (Z_1 > z_1, \ldots, Z_m > z_m). \)

Theorem 3.2 gives the posterior Laplace transform based on the complete observations combined with censored observations. Its proof follows directly from (3.2) and Theorem 3.1.

Theorem 3.2. The posterior Laplace transform of

\( h_1, \ldots, h_k \)

given complete and censored observations is

\[
\Psi(\xi_1, \ldots, \xi_k | X_1 = x_1, \delta_1, \ldots, X_n = x_n, \delta_n, Z_1 > z_1, \ldots, Z_m > z_m) = \Psi(\xi_1, \ldots, \xi_k | X_1 = x_1, \delta_1, \ldots, X_n = x_n, \delta_n) 
\cdot \exp \left\{ \sum_{j=1}^k \int T \int_{\mathbb{R}_+} (e^{-\varepsilon_j(s)} - 1) Q_j^*(dv, ds; z) \right\},
\]
where

\[ Q^*_j(dv, ds; z) = \exp \left\{ -v \sum_{j=1}^k \sum_{r=1}^m (z_r - s)^+ \right\} Q_j(dv, ds). \]

Now consider the problem of estimation of subsystem aggregate failure rates. In biological system, this corresponds to the problem of estimating the failure rate after some causes of failure have been removed from the environment. Recall that the aggregate failure rate of subsystem \( J \) is

\[ h_J(s) = \sum_{j \in J} h_j(s). \]

Suppose that the loss function is

\[ L(h_j, \hat{h}_j) = \int_T \left[ h_j(s) - \hat{h}_j(s) \right]^2 W(ds), \quad (3.5) \]

where \( W \) is a measure on \([0, \infty)\) that satisfies

\[ \int_T \var Y_j(s) W(ds) < \infty, \quad 1 \leq j \leq k. \]

In this case the Bayes estimate of \( h_j(t) \) is given by

\[ \hat{h}_j(t) = E(h_j(t)|X_1 = x_1, \delta_1, \ldots, X_n = x_n, \delta_m, Z_1 > z_1, \ldots, Z_m > z_m). \quad (3.6) \]

This Bayes estimate can be obtained from the posterior Laplace transform by using

\[ \xi_j(s) = \begin{cases} \Theta I(s < t), & j \in J, \\ 0, & j \notin J, \end{cases} \]

**Theorem 3.3.** If the loss function satisfies (3.5) then the Bayes estimate given both complete and censored observations is

\[
\hat{h}_j(t) = \sum_{i \in D^c \cap J} \frac{g_i(x(i, t); 0, \ldots, 0)}{g_i(x; 0, \ldots, 0)} + \sum_{i \in D \cap J} \int_0^t \int_{R^+} v Q_i(dv, ds; x) \nonumber \\
+ \sum_{i \in J} \int_0^t \int_{R^+} v Q^*_i(dv, ds; z), \nonumber
\]

where \( x_{ij}(i, t) = x_{ij}, \ r \neq i, \) and \( \{x_{ij}\}, \ 1 \leq j \leq n_i + 1, \) denotes the ordered values of \( x_{i1}, \ldots, x_{in}, \ t. \)

Note that the complete data contributes to the estimate \( h_j, j \in D^c \cap J, \) in the same way as the censored data. Also, this estimate can be thought of as a sequential estimate in the sense that the complete data represent systems that have failed before time \( t \) and the censored data represent systems that have not yet failed at time \( t. \)
4. Proportional hazards

The proportional hazards model of Cox (1972) represents the survival distribution of a device as

$$P(X > t | \Lambda) = \exp \{- \Lambda(t) e^{-\beta Z} \}, \quad (4.1)$$

where \( \Lambda \) is the hazard function of some baseline survival distribution, \( \beta \) is a vector parameters, and \( Z \) is a vector of constants which represents the measurement of some covariates for the device. Kalbfleisch (1978) and Burridge (1981) considered this model in a Bayesian context by regarding \( \Lambda \) as a realization of a gamma process. The goal is then to estimate \( \beta \) and \( \Lambda \) based on the failure times of a set of individuals all of which are exposed to the same realization of the gamma process.

The case in which \( \beta = 0 \) and \( \Lambda \) is a nonnegative process with independent increments has already been treated by Ferguson and Phadia (1979). Kalbfleisch and Burridge approach this problem when \( \beta \) is unknown by first obtaining an estimate of \( \beta \) from the marginal survival probability

$$E_{\Lambda} P(X > t) = E_{\Lambda} \exp \left\{ - \sum_{j=1}^{n} \Lambda(t_j) e^{-\beta Z} \right\}. \quad (4.2)$$

The (partial) Bayes estimate of \( \Lambda \) is then obtained by replacing \( \beta \) in (4.1) by its estimate \( \hat{\beta} \). As noted by Burridge, this model is inappropriate for the analysis of data recorded in continuous time since the gamma process is purely atomic.

The results of Section 2 provide a model for the case in which \( \beta = 0 \) and the survival distribution is absolutely continuous. In this section, it is shown how this model can be extended to treat and proportional hazards model for continuous data. Let \( Y_1 \in B \) and \( Y_2 \in B^* \) be independent random measures on \( T \) with Lévy measures \( Q_1 \) and \( Q_2 \) respectively. Next define the failure rate of the baseline survival distribution by

$$\lambda(t) = Y_1[0, t] + Y_2(t, \infty).$$

Then it can be seen that

$$P(X > t | \lambda) = 1 - F(t; \lambda, \beta) = \exp \left\{ -e^{-\beta \int_0^t \lambda(s) ds} \right\}. \quad (4.3)$$

Suppose that \( X_1, \ldots, X_n \) represent the failure times of \( n \) independent devices each of which experiences the same realization of \( \lambda \) and which have known covariates \( z_1, \ldots, z_m \) respectively. Let \( G(t) \) denote the marginal survival probability of \( X_1, \ldots, X_n \), defined in (4.4) below. Then \( G \) can be obtained directly from the joint Laplace transform of \( Y_1 \) and \( Y_2 \) and is given by

$$G(t) = P(X_1 > t, \ldots, X_n > t)$$

$$= E \exp \left\{ - \sum_{j=1}^{n} \exp(-z_j \beta) \int_0^t \lambda(s) ds \right\}.$$
\[ E \exp \left\{ \sum_{j=1}^{n} \exp(-z_j \beta) \left[ \int_T (t_j - s)^+ Y_1(ds) + \int_T (t_j \wedge s) Y_2(ds) \right] \right\} \]
\[ = \exp \left\{ \int_T \int_{\mathbb{R}_+} \left( \exp \left\{ -v \sum_{j=1}^{n} \exp(-z_j \beta)(t_j - s)^+ \right\} - 1 \right) Q_1(dv, ds) \right\} \]
\[ + \int_T \int_{\mathbb{R}_+} \left( \exp \left\{ -v \sum_{j=1}^{n} \exp(-z_j \beta)(t_j \wedge s) \right\} - 1 \right) Q_2(dv, ds) \right\}, \quad (4.4) \]

where the interchange of integration in (4.4) is justified by Tonelli's Theorem. Since this joint probability measure is absolutely continuous with respect to Lebesgue measure, a partial likelihood function for \( \beta \), as well as for any unknown parameters of the prior \( \text{Lévy measures} \ Q_1 \) and \( Q_2 \), can be obtained by differentiating (4.4).

For convenience, suppose that the observations are ordered, \( 0 \leq t_1 < t_2 < \cdots < t_n \). Using the notation of Section 2, define

\[ q(t; \beta, f, g) = \exp \left\{ -v \sum_{j=1}^{n} z_j \beta \right\} \sum_{\sigma \in \pi} \prod_{\tau \in \sigma} \int_{\mathbb{R}_+} I \left\{ s \leq \min(t_i) \right\} \]
\[ \cdot v^j e^{-v g(s)} Q_1(dv, ds ; t, \beta) \]
\[ + \int_{\mathbb{R}_+} \int_{\mathbb{R}_+} I \left\{ s > \max(t_i) \right\} v^j e^{-v g(s)} Q_2(dv, ds ; t, \beta) \right\}, \]

where

\[ Q_1(dv, ds ; t, \beta) = \exp \left\{ -v \sum_{j=1}^{n} \exp(-z_j \beta)(t_j - s)^+ \right\} Q_1(dv, ds), \]
\[ Q_2(dv, ds ; t, \beta) = \exp \left\{ -v \sum_{j=1}^{n} \exp(-z_j \beta)(t_j \wedge s) \right\} Q_2(dv, dr). \]

The likelihood function is given in Theorem 4.1 and is obtained as in Section 2.

**Theorem 4.1.** The likelihood function based on \( X_1, \ldots, X_n \) is given by

\[ \frac{\partial^n}{\partial t_1 \cdots \partial t_n} \exp\{G(t)\} = q(t; \beta, 0, 0) \exp\{G(t)\}. \quad (4.5) \]

If \( \beta \) is known or if \( \beta \) is estimated from the likelihood function (4.5), then it remains to find the Bayes estimate of \( \lambda \). The posterior joint Laplace transform of \( Y_1 \) and \( Y_2 \) can be obtained by applying Theorem 2.2.

**Theorem 4.2.** The posterior joint Laplace transform of \( Y_1 \) and \( Y_2 \) given \( X_1 = \)}
\[ t_1, \ldots, X_n = t_n, \text{ is} \]
\[
\Psi_t(f, g) = q(t; \beta, f, g)[q(t; \beta, 0, 0)]^{-1} \cdot \exp \left\{ \int_T \int_{\mathbb{R}_+} \left( e^{-g(s)} - 1 \right) Q_1(du, ds; t, \beta) + \int_T \int_{\mathbb{R}_+} \left( e^{-g(s)} - 1 \right) Q_2(du, ds; t, \beta) \right\}.
\]

There are two loss functions that will be considered for this problem,
\[
L_1(\lambda, \hat{\lambda}) = \int_T [\hat{\lambda}(t) - \lambda(t)]^2 W_1(dt),
\]
\[
L_2(\lambda, \hat{\lambda}) = \int_T [\hat{F}(t) - F(t)]^2 W_2(dt),
\]
where \( W_1 \) is a measure on \( T \) which satisfies
\[
\int_T \text{var} Y_1[0, t] W_1(dt) < \infty, \quad \int_T \text{var} Y_2(t, \infty) W_1(dt) < \infty,
\]
and \( W_2 \) is a finite measure on \( T \). For loss function \( L_1 \), the Bayes estimate is
\[
\hat{\lambda}(u) = E(\lambda(u)|X_1 = t_1, \ldots, X_n = t_n),
\]
and if \( L_2 \) is used, then
\[
\hat{F}(u) = 1 - E(\exp \left\{ -\int_0^u \lambda(t) dt \right\} |X_1 = t_1, \ldots, X_n = t_n).
\]
To obtain \( \hat{\lambda}(u) \), use
\[
f_1(s) = \Theta I\{s \leq u\}, \quad g_1(s) = \Theta I\{s > u\},
\]
in (4.6) and then differentiate with respect to \( \Theta \). To obtain \( \hat{F}(u) \), use
\[
f_2(s) = (u - s)^+, \quad g_2(s) = u \wedge s
\]
in (4.6).

**Theorem 4.3.** The Bayes estimates of \( \lambda \) and \( F \) are given by, respectively,
\[
\hat{\lambda}(u) = q(t(u); \beta(u), 0, 0)/q(t; \beta, 0, 0), \quad \hat{F}(u) = 1 - \Psi_t(f_2, g_2),
\]
where \( t(u) \) denotes the ordered values of \( u, t_1, \ldots, t_n \) and \( \beta(u) \) has 0 in the coordinate corresponding to \( u \) in \( t(u) \).

This section is concluded with a brief discussion regarding the inclusion of a 'strength of belief' parameter for this model. Suppose for example that the prior failure rate \( \lambda(t) \) is taken to be a gamma process with shape \( \alpha(t) \) and scale \( \Theta(t) \).
The Lévy measure of this process is

$$Q(\mu, dt) = \alpha(t) \mu^{-1} e^{-\mu/(\Theta(t))} \mu dt,$$

and

$$E\lambda(t) = \int_0^t \alpha(s) \Theta(s) \, ds, \quad \text{var} \lambda(t) = \int_0^t \alpha(s) \Theta^2(s) \, ds.$$

Note that this prior chooses an increasing failure rate with probability one. Now suppose, based on prior information, that the baseline failure rate is expected to be $\lambda_0(t)$, i.e., $E\lambda(t) = \lambda_0(t)$, where $\lambda_0$ is an increasing, differentiable function with nonzero derivative $\lambda_0'$. In this case $\eta(t) = \text{var} \lambda(t)$ provides a measure of the uncertainty or strength of belief regarding this choice of the failure rate. Suppose that $\eta$ is also differentiable with nonzero derivative $\eta'$. Hence, if $\lambda_0(t)$ and $\eta(t)$ are specified, then $\alpha$ and $\Theta$ become

$$\alpha(t) = [\lambda_0'(t)]^2 [\eta'(t)]^{-1}, \quad \Theta(t) = [\eta(t)][\lambda_0'(t)]^{-1}.$$

References


D.R. Cox, Regression models and life tables (with discussion), JRSS B 34 (1972) 187-208.


