ESTIMATION OF RANDOM SURVIVAL FUNCTION:
A LINEAR APPROACH
VICENTE QUESADA PALOMA, ALFONSO GARCÍA PÉREZ

In the first part of this work, a Survival function is considered which is supposed to be - an Exponential Gamma Process. The main statistical and probability properties of this process and its Bayesian interpretation are considered.

In the second part, the problem to estimate, from a Bayesian view point, the Survival function is considered, looking for the Bayes rule inside the set of linear combinations of a given set of sample functions.

We finish with an estimation, in the same situation like before, of the survival mean time, and the 1st moment about the origin of the Survival function.

Key words: Exponential Gamma Process, linear approach, Survival function, Bayesian nonparametric estimation, Survival mean time.

1. INTRODUCTION.

The application of the Bayesian method to -- the resolution of some statistical problems, has been made with enough success in parametric situations, but in nonparametric ones -- the application has not been so broad. The principal obstacle to do this has been to find workable priors on the set of all probability distributions on a given sample space.

/3/ and /2/, with Dirichlet processes and -- processes neutral to the right respectively, showed that these could be used to solve different nonparametric Bayesian decision problems.

/6/ presented an alternative method to the treatment of Bayesian nonparametric problems, approximating the solution of the original problem, when the decision space is restricted to be the set of linear combinations of some given set of sample functions.

Here we present a linear approach to the Survival function after to study the main statistical and probability properties, under the hypothesis the prior distribution over the space of probability distributions defined over a determined sample space, is an Exponential Gamma Process.

2. THE SURVIVAL FUNCTION.

Let $T$ be a nonnegative random variable, which we suppose indicates the failure time of a -- system or the time death of a live thing. The survival function of this random variable $T$ is defined like the probability of $T>t$, i.e.,

$$S(t) = \Pr \{ T>t \}, \quad vt>0$$

Let $F(t)$ be the distribution function of the random variable $T$, then

$$S(t) = 1 - F(t)$$

and so, the main properties of the Survival function will be,

a) $0 \leq S(t) \leq 1$.

b) $S(t)$ is a nonincreasing function.

c) $S(t)$ is a right-continuous function.

d) $\lim_{t \to 0} S(t) = 1$.

- Vicente Quesada Paloma, Departamento Estadística Matemática, E.I.O. Universidad Complutense de Madrid
- Alfonso García Pérez, Departamento Estadística Matemática, Universidad Autónoma de Madrid
- Artículo rebut el Desembre de 1981
The survival function can depend on one or more parameters; in this situation, we will write \( S(t; \theta) \) to mean a survival function conditional to the parameter \( \theta \).

In the parametric Bayesian situation, the parameter \( \theta \) is a random variable with prior distribution function \( G(\theta) \).

### 2.1. RANDOM SURVIVAL FUNCTION: DEFINITION.

A stochastic process \( S(t), t \in [0, \infty) \) is called a random survival function if

a) \( S(t) \) is a nonincreasing function, a.s.

b) \( S(t) \) is right-continuous in probability.

c) \( \lim_{t \to \infty} S(t) = 1 \) a.s.

d) \( \lim_{t \to \infty} S(t) = 0 \) a.s.

### 3. EXPONENTIAL GAMMA PROCESSES.

Let be the random survival function,

\[
S\left( t/\Lambda_0 \right) = \exp \left[-\Lambda_0(t) \right], \ t > 0
\]

which means \( S(t/\Lambda_0) = \Pr(T > t/\Lambda_0) \), where \( \Lambda_0(t) \) is a stochastic process which we will define through the definition of the random survival function. We have that,

\[
\lim_{t \to \infty} \Lambda_0(t) = 0 \quad \text{a.s., to be b) satisfied.}
\]

\[
\lim_{t \to \infty} \Lambda_0(t) = \infty \quad \text{a.s., to be c) satisfied.}
\]

\[
\Lambda_0(t) - \Lambda_0(s) \geq 0 \quad \text{a.s. if } t > s, \text{ to be a) satisfied.}
\]

Let be a partition of \( [0, \infty) \) in a finite number of disjointed intervals

\[
[0, a_1], (a_1, a_2], \ldots, (a_{k-1}, a_k = \infty).
\]

Let us see that the first interval is closed and last one open. If we write

\[
q_i = \Pr\left\{ T \in (a_{i-1}, a_i] / T > a_{i-1} \text{, } \Lambda_0 \right\}, \quad \text{if } \Pr\left\{ T > a_{i-1} \text{, } \Lambda_0 \right\} > 0 \quad \text{and } \quad q_i = 1 \text{ elsewhere, } \quad i = 1, \ldots, k
\]

3.1. PROPOSITION.

\[
\Lambda_0(a_i) = \sum_{j=1}^{i} -\log(1 - q_j) = \sum_{j=1}^{i} r_j,
\]

\( i = 1, \ldots, k \)

i.e., in each point \( a_i, i = 1, \ldots, k \), \( \Lambda_0(a_i) \) is the sum of i nonnegative variables.

Proof:

\[
q_i = \Pr\left\{ T \in (a_{i-1}, a_i] / T > a_{i-1} \text{, } \Lambda_0 \right\} = \frac{\Pr\left\{ T \in (a_{i-1}, a_i] \cap T > a_{i-1} \text{, } \Lambda_0 \right\}}{\Pr\left\{ T > a_{i-1} \text{, } \Lambda_0 \right\}} = \frac{\Pr\left\{ T \in (a_{i-1}, a_i] \text{, } \Lambda_0 \right\}}{\Pr\left\{ T = a_{i-1} \text{, } \Lambda_0 \right\}} = \frac{S(a_{i-1} - a_i) - S(a_{i-1})}{S(a_{i-1}) - S(a_{i-1} - a_i)}
\]

and if we change the survival function for its value, then

\[
\exp\left[-\Lambda_0(a_i)\right] = \prod_{j=1}^{i} (1 - q_j)
\]

so

\[
\Lambda_0(a_i) = \sum_{j=1}^{i} r_j
\]

where \( r_j = -\log(1-q_j), j = 1, \ldots, k \).

To specify a distribution over the trajectory space of the process \( S(t/\Lambda_0) \) is, by hand, the same that to specify a distribution over the trajectory space of the process \( \Lambda_0(t) \) and in the other hand, because the proposition proved, the same that to specify the finite dimensional distribution of the random variables \( r_1, \ldots, r_k \) over each partition \( (a_{i-1}, a_i] \) i = 1, ... k such that some consistence condition are satisfied.

So the problem is to specify a process, which according to the before proposition must be with independent increments and nondecreasing and in the other hand the distributions of the random variables \( r_i \) must be the same that the got one by apllication of the rules to the combined interval \( (a_{i-1}, a_{i+1}] \) .
2.2. Definition.

Let \( c > 0 \) and \( \Lambda^\theta(t) \) be a function of \( t \), such that \( \exp[-\Lambda^\theta(t)] \) is a nonrandom Survival function. We say that the random Survival function

\[
S(t/\Lambda^\theta_0) = \exp[-\Lambda^\theta_0(t)], \quad t \geq 0
\]

is an Exponential Gamma process with parameters \( (\Lambda^\theta(t), c) \), if for each finite partition \( (a_{i-1}, a_i] \), \( i = 1, \ldots, k \) of \([0, \infty)\), the random variables \( r_1, \ldots, r_k \) (the increments of the process) we saw before, are distributed independently like gamma with parameters \((c(\Lambda^\theta(a_i) - \Lambda^\theta(a_{i-1})), c)\), \( i = 1, \ldots, k \). We mean,

\[
r_k \equiv G(c(\Lambda^\theta(a_k) - \Lambda^\theta(a_{k-1})), c)
\]

\( \Lambda^\theta(t) \) is called shape parameter and \( c \) scale parameter.

3.3. Proposition.

Let \( S(t/\Lambda^\theta_0) \) be an Exponential Gamma Process with parameters \( (\Lambda^\theta(t), c) \). Then there exists a probability measure over the trajectory space of the \( \Lambda^\theta_0(t) \) process, and so over the trajectory space of the \( S(t/\Lambda^\theta_0) \) process, so over the parametric space \( \mathcal{T} \).

Proof:

\[
r_i \equiv G(c(\Lambda^\theta(a_i) - \Lambda^\theta(a_{i-1})), c), \quad i = 1, \ldots, k
\]

for each partition \( (a_{i-1}, a_i] \) and because the gamma distribution is reproductive in the first parameter,

\[
r_1 + r_{i+1} \equiv G(c(\Lambda^\theta(a_{i+1}) - \Lambda^\theta(a_i)), c)
\]

i.e., the same if we consider the interval \( (a_{i-1}, a_i] \), and so the consistency conditions are satisfied and the proposition is shown.


Let \( S(t/\Lambda^\theta_0) \) be an Exponential Gamma Process with parameters \( (\Lambda^\theta(t), c) \), then the mean of this process is \( (c/(c+1)) \Lambda^\theta(t) \cdot c \), i.e.,

\[
E[S(t/\Lambda^\theta_0)] = E[\exp(-\Lambda^\theta_0(t))] = (c/(c+1)) \Lambda^\theta(t) \cdot c
\]

where the expectation is taken with respect \( \mathcal{T} \), the prior probability over \( \mathcal{T} \).

Proof:

Let be the partition \([0, t), (t, \infty)\). In this situation,

\[
\Lambda^\theta_0(t) \equiv G(c(\Lambda^\theta(t), c)
\]

and so,

\[
E[\exp(-\Lambda^\theta_0(t))] = \left(\frac{c}{c+1}\right)^{\Lambda^\theta(t)} \cdot c = S_0(t)
\]

where the last equality is taken like notation.

4.2. Proposition.

The function \( S_0(t) = E[S(t/\Lambda^\theta_0)] \) is nonrandom Survival function. The proof of the proposition is easy and we only need to verify the properties of a nonrandom Survival function.

Because \( S_0(t) \) is a nonrandom Survival function, in a Bayesian situation, it can be considered like the "prior" Survival function.

In another way, the parameter \( c \) of a process can be considered like one that show the belief in our prior knowledge. So, when \( c \to \infty \) then \( \text{Var}(\Lambda^\theta_0(t)) \to 0 \), i.e., the random variable \( \Lambda^\theta_0(t) \) for each \( t \) is degenerate in the real number \( \Lambda^\theta(t) \) and the random Survival function \( S(t/\Lambda^\theta_0) \) is the nonrandom Survival function \( S_0(t) \). Our prior knowledge is maximum.

If \( c \to 0 \), then \( \text{Var}(\Lambda^\theta_0(t)) \to \infty \), and the posterior information over \( S(t/\Lambda^\theta_0) \) will null.

The other parameter \( \Lambda^\theta(t) \) shows us the prior knowledge.

4.3. Proposition.

Let \( S(t/\Lambda^\theta_0) \) be an Exponential Gamma Process with parameters \( (\Lambda^\theta(t), c) \) and with prior --
Survival function $S_\Lambda(t)$. Then if $x>y$

$$E_S(y) = \left[ \frac{c}{(c+1)} \right] \frac{c\Lambda^\delta(x)}{(c+1)/(c+2)} = S_\Lambda(x) \left[ \frac{S_\Lambda(y)}{1(c)} \right]$$

where $1(c) = \log((c+2)/(c+1))/\log((c+1)/c)$

Proof:

$$E_S(x) = E_S(y) = \left[ \exp \left( - \left( \Lambda_\Lambda(x) + \Lambda_\Lambda(y) \right) \right) \right]$$

and if we call $Y_1 = \Lambda_\Lambda(x) + \Lambda_\Lambda(y)$, $Y_2 = \Lambda_\Lambda(y)$

we are able to calculate the density function of the vector $(Y_1, Y_2)$ because the independent increment property of the process $\Lambda_\Lambda$ and so

$$E_S(x) = \left( \frac{c}{c+1} \right) \frac{c\Lambda^\delta(x)}{(c+1)} \frac{c\Lambda^\delta(y)}{(c+2)}$$

4.4. COVARIANCE FUNCTION OF THE EXPONENTIAL GAMMA PROCESS.

The covariance function of the Exponential Gamma Process with parameters $(\Lambda_\Lambda(t), c)$, is

$$r(x,y) = \left[ E_S(x, y) - E_S(x) \cdot E_S(y) \right]$$

$$= \left[ \frac{c\Lambda^\delta(x)}{(c+1)} \right] \left[ \frac{c\Lambda^\delta(y)}{(c+2)} \right]$$

and so, because the proposition 4.3,

$$r(x,y) = \left\{ \begin{array}{ll}
\left( \frac{c}{c+1} \right) \frac{c\Lambda^\delta(x)}{(c+1)} & \text{if } x > y \\
\left( \frac{c}{c+1} \right) \frac{c\Lambda^\delta(y)}{(c+1)} & \text{if } x < y
\end{array} \right.$$
THEOREM: Let \( S(t/A) \) be an Exponential Gamma Process with parameters \((\lambda(t), c)\) and prior distribution function \( F_0(t) = 1 - S_0(t) \). Let

\[
g(t): [0, \infty) \to \mathbb{R} \text{ a measurable function of the random variable } T, \text{ such that}
\]

\[
\int_T \left( \int_0^\infty g(t) \, dF(t) \right) \, d\mathcal{P}(T) < \infty \quad (4.7.1)
\]

Then, \( \int_T \left( \int_0^\infty g(t) \, dF(t) \right) \, d\mathcal{P}(T) \) is the expectation of \( g(T) \) respect the distribution \( F(T) \).

Proof:

Because (4.7.1) we can use Fubini's theorem and get result after to see that \( F_0(t) \) is -- the parameter of the Process neutral to the right \( S(t/A) \).

THEOREM: Let \( S(t/A) \) be an Exponential Gamma Process with parameters \((\lambda(t), c)\) and \( \lambda(t) \) continuous function of \( t \).

Let \( q_1: [0, \infty) \to \mathbb{R}^+ \cup \{0\} \) and

\[
q_2: [0, \infty) \to \mathbb{R}^+ \cup \{0\} \text{ two measurable functions, nonnegatives and such that for each } n \in \mathbb{N}
\]

\[\int_0^n q_1(s) \, ds < \infty \text{ and } \int_0^n q_2(s) \, ds < \infty \]

Let us suppose that

\[
\int_{\mathcal{F}} \left( \int_0^\infty q_1(t) \, S(t) \, dt \left( \int_0^\infty q_2(u) \, S(u) \, du \right) \right) \, d\mathcal{P}(S),
\]

\[1 = 1, 2, j=1, 2.\]

Then,

\[
\int_{\mathcal{F}} \left( \int_0^\infty q_1(t) \, S(t) \, dt \left( \int_0^\infty q_2(u) \, S(u) \, du \right) \right) \, d\mathcal{P}(S) =
\]

\[= \int_0^\infty q_1(t) \, S_0(t) \left( \int_0^t q_2(u) \, S_0(u) \, du \right) \, c(\mathcal{F}) \, dt
\]

\[+ \int_0^\infty q_1(t) \left( S_0(t) \right) \, \int_0^\infty q_2(u) \, S_0(u) \, du \, c(\mathcal{F}) \, dt,
\]

where \( S_0(t) \) is the prior Survival function.

Proof:

Because \( \lambda(t) \) is continuous, V.Quesada and A.Garcia Pérez (1.981),

\[
\int_{\mathcal{F}} \left( \int_0^\infty q_1(t) \, S(t) \, dt \left( \int_0^\infty q_2(u) \, S(u) \, du \right) \right) \, d\mathcal{P}(S) =
\]

\[= \int_0^\infty q_1(t) \, q_2(u) \, E \left[ S(t) \, S(u) \right] \, dt \, du\]

and we get the result if we see the proposition 4.3.

5. PROBLEMS IN NONPARAMETRIC BAYESIAN ESTIMATION.

Let \( X \) be a random variable which takes real values, with random distribution \( F \). Given a simple random sample \((X_1, \ldots, X_n)\) of \( X \), we want to make estimations of a function \( g(F) \), where \( g \) is a function over the space \( \mathcal{F} \), space of all probability distributions over \((\mathbb{R}, \mathcal{B})\).

Let us suppose that over the measurable space \((\mathcal{F}, \mathcal{G})\), where \( \mathcal{G} \) is a \( \sigma \)-field with respect to which \( g \) is measurable, there exists a probability measure \( \mathcal{P} \).

If we want to make inference about \( g(F) \), working with a.s.r.s. \((X_1, \ldots, X_n)\) using quadratic loss, we get the Bayes rule,

\[
g(X_1, \ldots, X_n) = \int_{\mathcal{F}} g(F) \, dS_1, \ldots, S_n(F)
\]

if this exists, but except \( \mathcal{P} \) would be the induced by a Dirichlet process in \((\mathbb{R}, \mathcal{B})\) (Ferguson, 1.973), the posterior probability is unhandle.

In this communication, we will find Bayes rules when,

a) \( \mathcal{P} \) is induced by an Exponential Gamma Process.

b) The loss function is quadratic.

c) We look for the Bayes rule inside the set of linear combinations of some given set of sample function.

6. ESTIMATION OF THE SURVIVAL FUNCTION.

Let \( S(t/\lambda_0) \), \( t > 0 \), be an Exponential Gamma Process with parameters \((\lambda(t), c)\), and let us make the next classical Bayesian analysis: we select a random sample of size \( n_1 \) and we find the Bayes rule \( S_1(t) \), with quadratic loss, for \( S(t/\lambda_0) \) inside the set of decision rules like

\[
a_t S_1(t) + b_t S_0(t)
\]

where \( S_0(t) \) is the prior Survival function and \( S_1(t) \) the empirical Survival function.
Then we select another sample, of size $n_2$ -- and again we look for the Bayes rule inside
the set of decision rules like

$$a_t S_{n_2}(t) + b_t \hat{S}_{1}(t)$$

We continue with the process taking samples
of sizes $n_3, n_4, \ldots, n_k$, finding the respective
Bayes rules $\hat{S}_3(t), \hat{S}_4(t), \ldots, \hat{S}_k(t)$, looking
respectively inside the set of decision rules like

$$a_t S_{n_i}(t) + b_t \hat{S}_{i-1}(t), \quad i=3, \ldots, k$$

6.1. THEOREM.

In the process we described before, the Bayes rule for the random Survival function $S(t/L_0)$ is

$$\hat{S}_k(t) = \frac{n_k}{(n_k-1)} \left( \frac{S_0(t)}{S_0(t)} \right)^{1(c)} \frac{S_0(t)}{-n_k S_0(t)+1} S_n(t) + \frac{1 - \left( \frac{S_0(t)}{S_0(t)} \right)^{1(c)}}{(n_k-1) \left( \frac{S_0(t)}{S_0(t)} \right)^{1(c)} - n_k S_0(t)+1} S_0(t)$$

and the Bayes risk is

$$R_{\min} = \frac{S_0(t)^{1(c)+1} - \left( S_0(t)^{21(c)+1} + \left( S_0(t)^{1(c)+2} - \left( S_0(t) \right)^2 \right) \right)}{1 + (n_k-1) \left( S_0(t) \right)^{1(c)} - n_k S_0(t)}$$

Proof:

We get the result if we make the Bayes risk

$$\int_\mathcal{F} \left( \int_\mathcal{S} \left( S(t) - a_t S_{n_1}(t) - b_t S_{i-1}(t) \right)^2 dQ(S_{n_1}(t)) \right) d\mathcal{F}(S), \quad i=1, \ldots, k$$

minimum, where $Q$ is the distribution of the random variable $S_{n_1}(t)$, and noting that ---

$$E \left[ S(t/L_0) \right] = S_0(t) \text{ and that}$$

$$E \left[ S^2(t/L_0) \right] = \left( S_0(t) \right)^{1(c)+1}.$$  

We see that is unnecessary to make several samples and that is enough to get only one - of size $n$.

If we call $p_n(t)$ the coefficient of $S_0(t)$, - then the coefficient of $S_n(t)$ is $1-p_n(t)$.

6.2. THEOREM.

The Bayes rule before obtained for $S(t/L_0)$ - is such that, for every fixed $t$,

$$S_n(t) \xrightarrow{a.s.} S(t/L_0) \quad \text{as } n \to \infty.$$  

The proof is easy because $p_n(t) \to 0$ as $n \to \infty$.

A similar Bayes rule can be obtained if we - consider censored data, using instead the empirical Survival function, the estimate of - Kaplan-Meier(1.958).

6.3. THEOREM.

The Bayes rule $\hat{S}_n(t)$ before obtained for $S(t/L_0)$ is such that,

$$\lim_{c \to 0} S_n(t) = S_0(t) \quad \text{and},$$

$$\lim_{c \to 0} S_n(t) = \exp \left[ - \lambda^4(t) \right]$$

7. THE SURVIVAL MEAN TIME ESTIMATION.

Let $F(t)$ be a random distribution function - of the random variable $T > 0$, and let $S(t/L_0)$ the random Survival function associated which is an Exponential Gamma Process with prior Survival function $E \left[ S(t/L_0) \right] = S_0(t)$ known and with all the moments.

We shall call survival mean time, if it is exits to

$$\mu(F) = \int_0^\infty dF(t) = - \int_0^\infty dS(t) =$$

$$= - S(t,t) + \int_0^\infty S(t) dt = \int_0^\infty S(t) dt = \mu(S)$$

Qustistio - V. 6, n° 2 (june 1982)
7.1. **Problem Without Sample.**

The Bayes rule with respect to quadratic loss, will be the expected value of \( \mu(S) \) with respect to the distribution \( \mathcal{F} \).

7.2. **Estimation with Samples of Size \( n \).**

If we look for the Bayes rule inside the set of decision rules that are linear combinations of the sample mean \( \bar{X} \) and the \( \mu_0 \),

\[
\hat{\mu}(S) = a \mu_0 + b \bar{X}
\]  
(7.2.1)

we have the next result,

**Theorem:** The values of \( a \) and \( b \) that they do minimum (7.2.1), are

\[
b = \frac{n(1 - \mu_0)}{\mu_0^2 + (n-1)K_{11} - \mu_0} \quad , \quad a = 1 - b
\]

where,

\[
\mu_0 = \int_0^\infty (x)^2 dF_0(x) = -\int_0^\infty (x)^2 dG_0(x)
\]

and,

\[
K_{11} = \int_0^\infty \frac{c}{c+1} \left( \mu^2 - \mu_0^2 \right) \left( \frac{c}{c+2} \right) cA^+(x) dy dx + \int_0^\infty \frac{c}{c+1} \left( \mu_0^2 - \mu^2 \right) \left( \frac{c}{c+2} \right) cA^+(x) dy dx
\]

**Corollary:** If \( n \rightarrow \infty \), then the estimate \( \hat{\mu}(S) = \bar{X} \), i.e., the prior knowledge disappears and the only information is the sample one.

**Corollary:** The Bayes risk we reach, which -- goes to naught when \( n \) goes to naught when \( n \) goes to , is

\[
P_{\text{min}} = \frac{K_{11} - (\mu_0)^2}{\mu_0^2 + (n-1)K_{11} - n(\mu_0)^2}
\]

8. **The \( i \)th Moment Estimation.**

In the same situation like before, if we want to find the Bayes rule for the \( i \)th moment about the origin of \( F(t) \), we have the next theorem,
where

\[
K_{hj} = \sum_{j=1}^{J} \int_{0}^{\infty} \int_{0}^{\infty} e^{-x} y^{j-1} \left( \frac{c}{c+1} \right) \left( c+1 \right) \left( c+1 \right) c \Lambda(x) \left( \frac{c+1}{c+2} \right) c \Lambda(y) \left( \frac{c+1}{c+2} \right) c \Lambda(x) \left( \frac{c+1}{c+2} \right) c \Lambda(y) \left( \frac{c+1}{c+2} \right) dy \, dx
\]

9. REFERENCES.


