Kaplan Meier's Bayesian model under an informative prior distribution Case: integration study of unemployed registered with the Local Employment Agency of Ain El Benian (January 2011-July 2013)

*** Ahmed Hamimes [1], Benamirouche Rachid[2]**

[1]École nationale de la statistique et d'économie appliquée, Algérie,ahmedhamimes@yahoo.com.
[2] École nationale de la statistique et d'économie appliquée, Algérie, rbena2002@hotmail.com

Received: 13/10/2020 *Accepted: 25/11/2020* *Published: 31/12/2020*

Abstract

A Bayesian approach to nonparametric survival offers practical, simple and relatively easysolutions to exploit numerically.

In this contribution, we will demonstrate the efficiency of the Bayesian approach in themodeling of durations and in an econometric context, we propose a new conception of theKaplan Meier Bayesian estimator under an a priori informative law based on the stochasticapproximation. Which here represents by Gibbs sampling.Our contribution is to improve thedeductive stage in estimating nonparametric survival times and under censorship, and this iswhat we reached in our research.

Keywords: Bayesian approach, nonparametric survival, Kaplan Meier Bayesian estimator,Gibbs sampling.

JEL classification code : C11, C15, C41, E24.

Résumé

L'approche bayésienne de la survie non paramétrique offre des solutions pratiques, simples et relativement faciles à exploiter numériquement. Dans cette contribution, nous démontrerons l'efficacité de l'approche de Bayes dans la modélisation de la durée et dans le contexte économétrique, en proposant un nouveau concept pour l'estimateur bayésien de Kaplan Meyer sous une loi informationnelle a priori basée sur l'approximation stochastique. Ce qui est représenté ici par l'échantillonnage de Gibbs. Notre contribution est d'améliorer la phase déductive dans l'estimation des temps de survie non paramétriques et sous censure, et c'est ce que nous avons trouvé dans nos recherches.

Mots clés : approche bayésienne, survie non paramétrique, estimateur bayésien de Kaplan-Meier, échantillonnage de Gibbs.

Code de classification JEL : C11, C15, C41, E24.

*** Corresponding author**.

I. Introduction

The study of censored lifetimes is used in various fields of research, and various possibilities f or modeling these data have been suggested. The first field of application of the duration data analysis was in the biomedical sciences where it is used either for therapeutic or epidemiologi cal trials.In economics, we study the length of time spent unemployed, in a job or between two jobs, the length of a transportation trip, a business' lifespan, or the length of a "revolving" type loan.

In Bayesian inference, the survival analysis has gained growing attention in recent years, but still restricted due to the scarcity of specialized tools (one of the reasons of this scarcity is the difficulty in automating Bayesian analyzes compared to the frequentist approach), as well as the strength of habit and the difficulty in embracing a particular statistical conception.Bayesian inference 's great interest in frequentist approaches is its great clarity and coherent methodology in the theoretical program , which allows richer and more straightforward explanations to deduce results than those offered by the classical approach.

The Kaplan-Meier approach appears more fitting and more accurate in the traditional nonparametric simulation of survival times in most of the conditions encountered today in clinical research. In some special cases the actuarial method may still retain indications for use today if the occurrence times of unknown events or the size of the study are high.In this contribution, we will give the actuarial estimator and classical Kaplan Meier a Bayesian alternative based on the beta distribution, after we present several methods for calculating the prior distribution.For many reasons, we find this necessary to provide a Bayesian estimator based on these two classical concepts, firstly because, for example, in medical science, very often we have to deal with small samples. There are several explanations for that. The most popular is the disease's severity, or the difficulty of getting patients with the same biochemical parameters together.Additionally, we have censored data quite often. Therefore, a small sample size does not allow us to use classical statistical methods, or they can give us very common and even false results when they are used. The second explanation we achieve results for both censored and uncensored cases, using the Bayesian method.That is why our survival curve is smoother, and for the chance of survival, we don't have such fast jumps. Furthermore the interval is even smaller for some observations.In the case of the Kaplan Meier or Actuarial estimator, resorting to the Bayesian method simplifies the use of this approach compared to other methods (methods based on randomized measurement), and addresses the issue of habit strength. Where frequentist calculations are used.Several works were based upon the Kaplan Meier estimator 's creation. Rossa and Zieliński (1999) used the Kaplan-Meier estimator 's local smoothing solution based on an approximation by the Weibull distribution function. Rossa and Zieliński (2002), using Weibull 's law as a Kaplan Meier process approximation function.Shafiq Mohammad et al (2007), provided a weighting of the Kaplan Meier estimator under the sine function for heavily censored results. Khizanovand Maĭboroda (2015) used a mixing model of varying concentrations to adjust or alter the Kaplan Meier estimator.

In this contribution, the classic Kaplan Meier estimator will give a Bayesian alternative based on the distribution of beta informative. The durations of global unemployment in Ain El Benian’s National Employment Agency (ANEM) will be analyzed in our application. Work carried out on a sample of 1064 unemployed people observed between 01/01/2011 and 15/07/2013.This application makes it possible to demonstrate in practice that Bayesian econometrics procedures constitute an essential element for the mastery of economic information due to the difference that exists in the interpretation and estimation of curves and durations of the exiting unemployment.

II. Modèle statistique Bayésien

We write that if an event E can be the result of a cause A or a cause B, according to Laplace's principle, the probability of one of the causes:

$$p(A/E) = \frac{P(E/A)}{P(E/A) + P(E/B)}$$

when$P(E) > 0$, according to the conditional probability rule, we write:

$$P(A_k/E) = \frac{P(A_k \cap E)}{P(E)}$$

When the events $A_1, A_2, \dots, A_n$,E satisfy the conditions of total probability:$P(E) = \sum_{i=1}^{n} P(A_i)\, P(E/A_k)$. By commutativity, we establish:

$$P(A_k \cap E) = P(A_k)P(E/A_k)$$

we find as a result, Bayes' law:

$$P(A_k/E) = \frac{P(A_k)P(E/A_k)}{\sum_{i=1}^{n} P(A_i)\, P(E/A_k)} \qquad (1)$$

Bayes published in 1763, after his death because he was not sure of his conclusions. Ficher generates by the use of the conditional probability an estimation method which consists in maximizing by the parameter this a posteriori probability, by using "modern" notations, with $f(x/\theta)$the a priori distribution of the rv$X$for a value of the parameter θ, $g\ (\theta)$ and the law of the parameter θ on the reference population, the marginal density of the sample$(x_1, \dots, x_n)$is:

$$f(x_1, \dots, x_n) = \int_{\Theta} \prod_{i=1}^{n} g(\theta) f(x_i/\theta)\, d\theta$$

Conditionally to$(x_1, \dots, x_n)$, the posterior distribution of θ is therefore:

$$\pi(\theta/x_1, \dots, x_n) = \frac{likelihood \times prior}{\int likelihood \times prior} = \frac{\prod_{i=1}^{n} f(x_i/\theta) \times g(\theta)}{\int_{\Theta} \prod_{i=1}^{n} f(x_i/\theta) \times g(\theta) d\theta} \ (2)$$

This a posteriori law is the combination of:

- $f(x/\theta)$the density function of x knowing the value of the random variable θ.
- $\pi(\theta)$models the a priori density function on θ.

- $m(x)$the marginal distribution of x.

Once the data are available, the quantity $m(x)$ in equation (2) is a normalization constant which guarantees that $\pi\ (\theta\ /\ x)$ is indeed a probability distribution. We can write:

$$\pi(\theta/x) \propto f(x/\theta) \times \pi(\theta)(3)$$

III. Méthode de Monte Carlo par Chaine de Markov

From a practical point of view, there are situations in which calculating the Bayesian estimate of magnitude of interest is particularly difficult by hand; these cases require simulation methods based on stochastic algorithms involving iterative sampling procedures.

Definition 1 (The Monte Carlo Markov Chains method (MCMC)).
The Monte Carlo method by Markov chains is any method, which produces an ergodic Markov chain whose stationary distribution is the distribution of interest.
The two most popular algorithms are the Metropolis-Hastings algorithm and the Gibbs sampling algorithm shown below.

- **Gibbs sampling method**

Gibbs sampling is the most widely used Bayesian algorithm in statistical inference, this method was used by Geman and Geman in 1984 to generate observations from a Gibbs distribution (Boltzmann distribution). Among the properties of this algorithm, we find:

- Gibbs sampling is a special case of the M-H algorithm such that the acceptance probability is always equal to 1;
- Gibbs sampling takes advantage of the hierarchical structures of a model;
- in some algorithms it is found that the simulations can be rejected, on the other hand the Gibbs sampling where all the simulations are taken into account;
- this algorithm is well suited to a model with hidden information, whether with censorship or with a latent variable.

The algorithm breaks down into the following points:

- Initialize $\theta_0 = \theta_0^{(1)}, \theta_0^{(2)}, \dots, \theta_0^{(k)}$, which is the first vector of elements in the string.
- Set $t \leftarrow 0$.
- To go from step t to step $t\ +1$:
  - Generate $\theta_{t+1}^{(1)}$ by simulating according to the law $\pi_1\left(\theta_t^{(1)} / \theta_t^{(2)}, \dots, \theta_t^{(i)}, \dots \theta_t^{(k)}\right)$

- Generate $\theta_{t+1}^{(2)}$ by simulating according to the law $\pi_2\left(\theta_t^{(2)}/\theta_{t+1}^{(1)},\theta_t^{(3)},\dots\theta_t^{(k)}\right)$
- ⋮
- Generate $\theta_{t+1}^{(k)}$ by simulating according to the law $\pi_k\left(\theta_t^{(k)}/\theta_{t+1}^{(1)},\theta_{t+1}^{(2)}\dots\theta_{t+1}^{(k-1)}\right)$

- Change the value of t to $t \leftarrow t+1$, and go to 3.

For more details see (Robert, C.P (2014), Robert, C.P. (2006), Robert, C.P., Casella, G. (2004)).

IV. Le critère de déviance d'information et le choix du modèle

Spiegelhalter et al 2002, propose a generalization of AIC and BIC[1] (Bayesian Information Criterion) because the asymptotic justification is not appropriate in hierarchical models. This criterion is more satisfactory than the previous ones because it takes into account the a priori information and integrates a penalization factor natural to the log-likelihood (Robert, 2006).

We assume the model $\mathfrak{M}$ with parameter θ, x the data sample, the information deviance criterion for $\mathfrak{M}$is written:

$$DIC_{\mathfrak{M}} = -4\, E[logf(x/\theta)/x] + 2logf(x/E(\theta/x))$$

The model selected is the one with the smallest DIC[2] value. Spiegelhalter et al (2002), propose that a difference of 5 or 10 in the value of DIC is generally negligible.

DIC has several desirable characteristics. First, the DIC is easy to calculate when the likelihood function is available in closed form and the posterior distributions of the models are obtained by Markov Chain Monte Carlo simulation (MCMC).

Second, it is applicable to a wide range of statistical models. Third, unlike Bayes factors (BF), it is not subject to the Jeffrey-Lindley paradox and can be calculated when uninformative or improper priors are used. Unlike the Bayes factor, the DIC size is not the objective: only the differences between the DICs are important.

V. Estimation de Kaplan Meier et l'approche Bayésienne

a. Estimateur de Kaplan-Meier[3]

The Kaplan-Meier (KM) estimation method is also called "Product Limit Estimations (PLE)" by Anglo-Saxon statisticians. This estimator, which generalizes the notion of empirical distribution function, is based on the following idea: living after a while means being alive just before and not dying over time, that is,

$$P(X > t) = P(X > t_{i-1}, X \geq t_i)$$

$$= P(X > t_i / X \geq t_{i-1})P(X > t_{i-1})$$

[1] The relevance of this criterion in a Bayesian context is questionable.

[2] The models compared all derive from the same model (Burnham and Anderson, 2002).

$$= P(X > t_i / X \geq t_{i-1}) P(X > t_{i-1} / X \geq t_{i-2}) P(X > t_{i-2})$$

Wecanwrite:
$S(t_i) = P(X > t_i / X \geq t_i) * S(t_{i-1})$

the proportion q_i of individuals who experienced the event at time t_i corresponds to

$$q_i = \frac{d_i}{n_i}$$

this quantity estimates the value of the risk function $\lambda(t)$for$t = t_i$, where :t_i represents the follow-up time since inclusion in the study for each patient$i$;

$d_i$est le nombre de décès au temps t_i;

$n_i$is the number of subjects at risk of presenting the event studied at the instant t_i, i.e. the number of patients who have not yet undergone the event nor the censorship just before t_i.

$(1 - q_i)$represents the proportion of people who did not experience the event.

The probability of survival in t_ide then becomes:
$$S(t_i) = S(t_{i-1})(1 - h_i)$$
$$= S(t_{i-1}) \frac{n_i - d_i}{n_i}$$

according to this equation, the probability of survival t_i knowing that we were alive at t_{i-1} is estimated as follows:

$$\hat{S}(t_i / t_{i-1}) = \frac{n_i - d_i}{n_i}$$

By extension, if we consider $t_1 < t_2 < \cdots < t_n$ the distinct survival times of n individuals, $\hat{S}(t)$corresponds to the product of all the probabilities of not having known the event since the start of the observation, but this estimator can be written in two different ways depending on whether there is no joint or the presence of a joint as follows:

- Case of absence of a joint (ex aequo):

L'estimateur de Kaplan-Meier est donné par

$$\hat{S}(t) = \begin{cases} \prod_{t_i \leq t} \left(1 - \frac{1}{n - i - 1}\right)^{\delta_i} & \text{if } t \geq t_1 \\ 1 & \text{if } t < t_1 \end{cases}$$

with

$$\delta_i = \begin{cases} 1 & \text{realized event,} \\ 0 & \text{censored subject.} \end{cases}$$

- Ex aequo presence:

In the case of application, we are confronted with the presence of events of different natures, we consider that the uncensored observations take place before the censored ones, we have

$$\hat{S}(t) = \begin{cases} \prod_{t_i \leq t} \left(1 - \frac{d_i}{n_i}\right) = \prod_{t_i \leq t} (1 - h_i) & \text{if } t \geq t_1 \\ 1 & \text{if } t < t_1 \end{cases}$$

The variance of the Kaplan-Meier estimator is given by Greenwood's formula:

$$Var[\hat{S}(t)] = \hat{S}(t)^2 \times \sum_{i|\ t_i \leq t} \frac{d_i}{(n_i - d_i)n_i}$$

The Kaplan-Meier estimator offers several advantages:

- It serves as a theoretical basis for any survival study because this type of estimator allows the estimation of one of the laws characterizing the duration variable without making a priori hypothesis on it.
- The Kaplan-Meier estimator is a maximum likelihood estimator.
- This estimator is a piecewise constant function, continuous to the right and limit to the left, with a jump at each observed death time.
- The Kaplan-Meier estimator is convergent, coherent and asymptotically Gaussian. In addition, it is positively biased.

There are several remarks about the Kaplan-Meier estimator:

- A Kaplan-Meier estimator can also be obtained for truncated data but not for interval-censored data (since times to death are not known).
- This estimator requires knowing exactly all the study entry and exit dates of all individuals; which is a drawback if the registers are poorly informed.
- This method is relevant if each time interval considered is small relative to the speed of the variation of the survival function. This is to ensure that the discretization does not generate a significant bias on the estimate.
- The Kaplan and Meier estimator cannot take into account the effect of individual characteristics by breaking down the study population into subpopulations (eg sex, PSC,..).
- If the number of observations in a study is large, and the survival time for each observation is measured precisely, the data collected will have too many distinct values. Therefore, the Kaplan and Meier estimator of the survival function generates a large data table, and a poorly represented graph.

b. La conception bayésienne de l'estimateur de KM

In the frequentist approach the number of deaths in the interval of time is an realization of a Binomial law written by:

$d_i \sim \beta in(n_i, q_i)$

where

$$q_i = 1 - \frac{d_i}{n_i} \quad (4)$$

For a binomial distribution and a conjugate prior distribution, we set

$$f_\pi(d_i/\alpha,\beta) = \int_0^1 f(d_i/q_i)\,\pi(q_i/\alpha,\beta)dq_i$$

$$= \int_0^1 [q_i(1-q_i)]^{-1} C_{n_i}^{d_i} q_i^{d_i}(1-q_i)^{n_i-d_i}\,dq_i$$

$$= C_{n_i}^{d_i}\frac{1}{B(\alpha,\beta)}\int_0^1 q_i^{d_i+\alpha-1}(1-q_i)^{n_i-d_i+\beta-1}\,dq_i$$

$$= C_{n_i}^{d_i}\frac{B(\alpha+d_i,n_i+\beta-d_i)}{B(\alpha,\beta)}$$

which provides a beta – binomial distribution to estimate $\hat{\alpha},\hat{\beta}$, in order to calculate $\pi(q_i/d_i,\hat{\alpha},\hat{\beta})$.

also let :

$$\begin{cases} n_1 = le\ nombre\ de\ sujet\ dans\ le\ début\ d\ l'étude \\ n_i = n_{i-1} - d_i - c_i \end{cases},$$

• Under the vague prior

A vague a priori law, it is a proper law with a very large variance, according to this distribution, the a priori law is considered as being weak informative, and one uses this law for the regularization and the stabilization, it provides solutions in the use of algorithms. We ask:

$$q_i \sim \beta(0{,}01{,}0{,}01)$$

• Under Jeffreys' a priori law

The a priori measure of Jeffreys (possibly improper) defined by:

$$\pi^*(q_i) \propto I^{1/2}(q_i)$$

we assume $d_i \sim \beta in(n_i; q_i)$, the probability distribution of d_i and the Fisher information are respectively

$$P(d_i = j) = C_{n_i}^{j} q_i^{\,j}(1-q_i)^{n_i-j}\ ; j \in \mathfrak{X} = \{0,1,2,..n\}, 0 < q_i < 1$$

$$I(q_i) = -E\left(\frac{\partial^2 lnf(j/q_i)}{\partial q_i^{\,2}}\right) = \frac{E(j)}{q_i^{\,2}} + \frac{E(n-j)}{(1-q_i)^2} = \frac{n_i}{q_i} + \frac{n_i}{1-q_i} = \frac{n_i}{q_i(1-q_i)}$$

Jeffreys' prior law is:

$$\pi^*(q_i) \propto I^{1/2}(q_i) \Rightarrow \pi^*(q_i) \propto n^{1/2}\left(q_i(1-q_i)\right)^{-1/2}$$

if we set $\pi(q_i) = A * \pi^*(q_i)$, and we integrate the two parts of the equation with respect to q_i, we find

$$1 = \int_0^1 A * n_i^{1/2} \left(q_i(1-q_i)\right)^{-1/2} dq_i$$

$$= A * n_i^{1/2} \int_0^1 \left(q_i(1-q_i)\right)^{-1/2} dq_i$$

$$= A \times n_i^{1/2} \times \beta(1/2;1/2)$$

so we have :

$$A = 1/\left(n_i^{1/2} \times \beta(1/2;1/2)\right)$$

starting from A, the a priori function is written as follows:

$$\pi(q_i) = A * \pi^*(q_i) = \frac{1}{\sqrt{n}\beta(1/2;1/2)} n^{1/2}\left(q_i(1-q_i)\right)^{1/2-1}$$

$$= \frac{1}{\beta(1/2;1/2)} q_i^{-1/2}(1-q_i)^{-1/2}; 0 < q_i < 1$$

so

$$q_i \sim \beta\left(\frac{1}{2}, \frac{1}{2}\right)$$

Jeffrey's law is criticized by some Bayesians as being a tool without subjective justification in terms of a priori information, in what follows we introduce two other distributions.

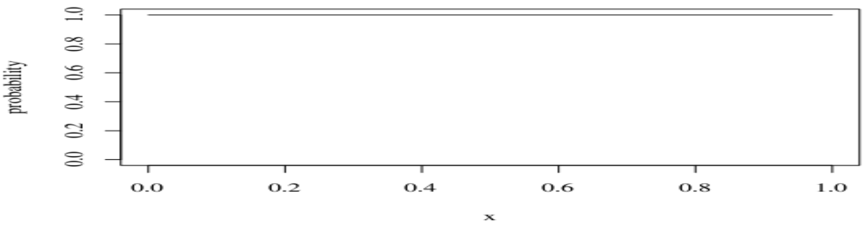
- **Under the a priori uniform law**

A first natural idea in order not to influence the results a priori is to consider each case as equiprobable and therefore to take a uniform measurement with respect to the Lebesgue measure or the counting measure in the discrete case.

The beta law which gives different representations offers the possibility of an a priori uniform modeling under the following law:

$$q_i \sim \beta(1,1)$$

Fig 1: the distribution of beta (1; 1)



Source: Developed by us.

- **The hierarchicallinkfunction**

For the law of recurrences $q_i \sim beta(\alpha, \beta)$ we assume a hierarchical construction, when $q_i$it is a probability belongs to the interval $[0,1]$, to increase the precision of the method estimate we want to link it to a series of regressors (exogenous variables), so we have to send it to IR.

One of the methods to introduce the relation between the probability $q_i$and the independent and unknown causal series is the logistic transformation given by:

$$logit(q_i) \overset{\text{def}}{=} ln\frac{q_i}{1-q_i}, \qquad q_i \in]0,1[$$

and

$\mu_i = logit(q_i)$, i.e.

$$q_i = \frac{exp(\mu_i)}{1+exp(\mu_i)},$$

When we introduce the information that in reality we have for each individual several observations at different times, it becomes necessary to resort to hierarchical Bayesian methods, thus a hierarchical Bayesian model is a compromise between the noninformative laws of Jeffreys , which are diffuse but sometimes difficult to use and explain, and the combined laws, which are subjectively difficult to justify but numerically practical.
Our problem remains nonparametric, we pose a Gaussian model for μ_i with unknown hyperparameters as follows:

$$\mu_i \sim \mathcal{N}(\vartheta;\tau)$$

$$\vartheta \sim \mathcal{N}(0;0{,}001)$$

This last proposition shows the principle of exchangeability, which reflects the conditional independence of the parameters.

In the a priori modelization of the dispersion, we use distributions with heavier tails, in the latter we find particularly two statistical distributions, Gamma and Cauchy.

In our case the hyperparameter of the variance τ is only made up of positive values, so we set the demi-Cauchy law. τfollows a half-Cauchy law if that density is:

$$\pi(\tau) = 2/\pi\rho\,[1+(\tau/\rho)^{-1}], \qquad \tau > 0$$

on note

$$\tau \sim HC(\rho)$$

$\rho$is the median of demi-Cauchy, which means that a priori beliefs are easily expressed.

The proposed model is written as follows

$$q_i = \frac{exp(\mu_i)}{1+exp(\mu_i)}$$

$$\mu_i \sim \mathcal{N}(\vartheta;\tau)$$

$$\vartheta \sim \mathcal{N}(0;0{,}001),\ \tau \sim HC(B)$$

$$B \sim Uniforme(0;\mathcal{T}),\ \text{we pose}: \mathcal{T} = 100.$$

i. The informative a priori distribution

Informative a priori laws represent the subjective way of thinking where the a priori is based on the information available on the parameter obtained. Among these methods we find the histogram approach, the relative likelihood and the conjugate approach remains the most standard solution in the informative framework. In another way, a measure π is informative if its relative[3] entropy (or differential) is finite by:

$$\int_{\Theta} \pi(\theta) log\pi(\theta) d\theta < +\infty,$$

However, it is possible to construct a priori laws by knowing the results of random experiment governed by a series of random variables, following a distribution $f\ (x\ /\ \theta)$, we will say that the prior law is objective, like a technical tool. When no a priori information is available, these a priori laws are considered non-informative[4].

From a Bayesian perspective we assume an a priori for$q_{i,}$, and when the distribution used in the informative case for the proportions is that of Beta, we set:

$q_{i\sim} beta(\alpha,\beta)$

VI. Application in econometrics of duration models

a. Data presentation

The data available relate to a filtered sample of 1064 unemployed registered with the local employment agency of Ain El Benian, over the period from January 01, 2011 to July 15, 2013. Distinguishing those who have found a job. employment, the placement of the unemployed during this period gives rise to 875 right-censored observations. In this case, the variable i represents the indication that the ith unemployed person entered a job after his daily period of unemployment t_i.

According to the information available on the unemployment rate in Algeria stood at 11.4% in May 2019, according to official figures published by the National Statistics Office (ONS) and released this Sunday, December 29 via the official agency. In total, the official number of the unemployed population was estimated at 1.449 million in May 2019.

So it is possible to construct several a priori in the mean is $\alpha\ /\ ((\alpha + \beta))$.As a sample of these laws we propose the following laws:

$\mathcal{L}_1 : q_i^1 \sim beta(0.1,0.77)$

$\mathcal{L}_2 : q_i^2 beta(1,7.77)$

$\mathcal{L}_3 : q_i^3 beta(3,23.31)$

[3]Relative entropy is an indicator of the uncertainty of a measurement, such as variance, but the latter may not be defined. We can therefore perceive an informative measure as possessing a measurable (and therefore calibratable) uncertainty indicator.

[4]En réalité, les lois a priori non informatives n'existent pas vraiment, et tous les a priori sont informatifs en quelque sorte. En effet, on connait au moins son domaine de variation, c'est-à-dire l'ensemble des états de la nature, et le rôle de chaque composante du paramètre sur les observables (paramètre de localisation, d'échelle, etc.).

$\mathcal{L}_4: q_i^4 beta(5{,}38.85)$

$\mathcal{L}_5: q_i^5 beta(6{,}46.63)$

$\mathcal{L}_6: q_i^6 beta(12{,}93.26)$

To select the appropriate model, we use the information deviance criterion (DIC) as follows:

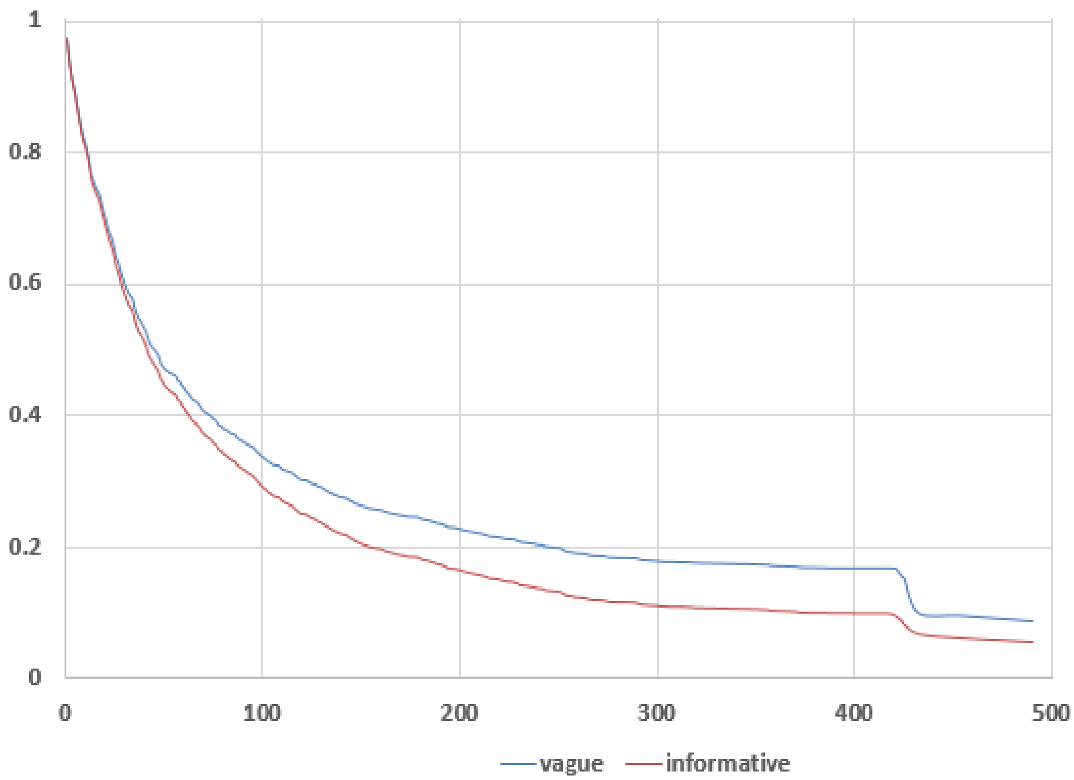
Tab 1 :the information deviance criterion for the proposed models.

models	**DIC**
$\mathcal{L}_1: q_i^1 \sim beta(0.1, 0.77)$	1117
$\mathcal{L}_2: q_i^2 beta(1{,}7.77)$	**1060**
$\mathcal{L}_3: q_i^3 beta(3{,}23.31)$	1292
$\mathcal{L}_4: q_i^4 beta(5{,}38.85)$	1646
$\mathcal{L}_5: q_i^5 beta(6{,}46.63)$	1835
$\mathcal{L}_6: q_i^6 beta(12{,}93.26)$	2953

Source: Developed by us.

From Table (1) we can clearly see that the best model that minimizes the information deviance criterion is the model with the second prior distribution.

Fig2: The Kaplan-Meier survival functions under the vague (beta (0.01, 0.01)) and informative prior law for the overall unemployment duration.



Source: Developed by us.

From Figure (2), we notice that at the start of the curve, 100% of the individuals in the sample are unemployed. After approximately 2 months of registering with this agency, 50% of individuals were placed in the labor market. But, the exit from unemployment for the rest of the individuals in the sample is spread out over a long period, for some it even exceeds a year. In general, from the unemployment duration curve, we deduce that the probability of leaving unemployment for those registered with the Local Employment Agency of Ain el Benian becomes very low for an unemployed person who exceeds more than a year of unemployment. This figure (2) also informs us that the median duration of exit from unemployment is 45 days in both methods. Thus, the two Bayesian estimation methods(vague and informative) are identical approximations although the method based on the informative prior distribution presents a preferable deviance criterion as the following table indicates:

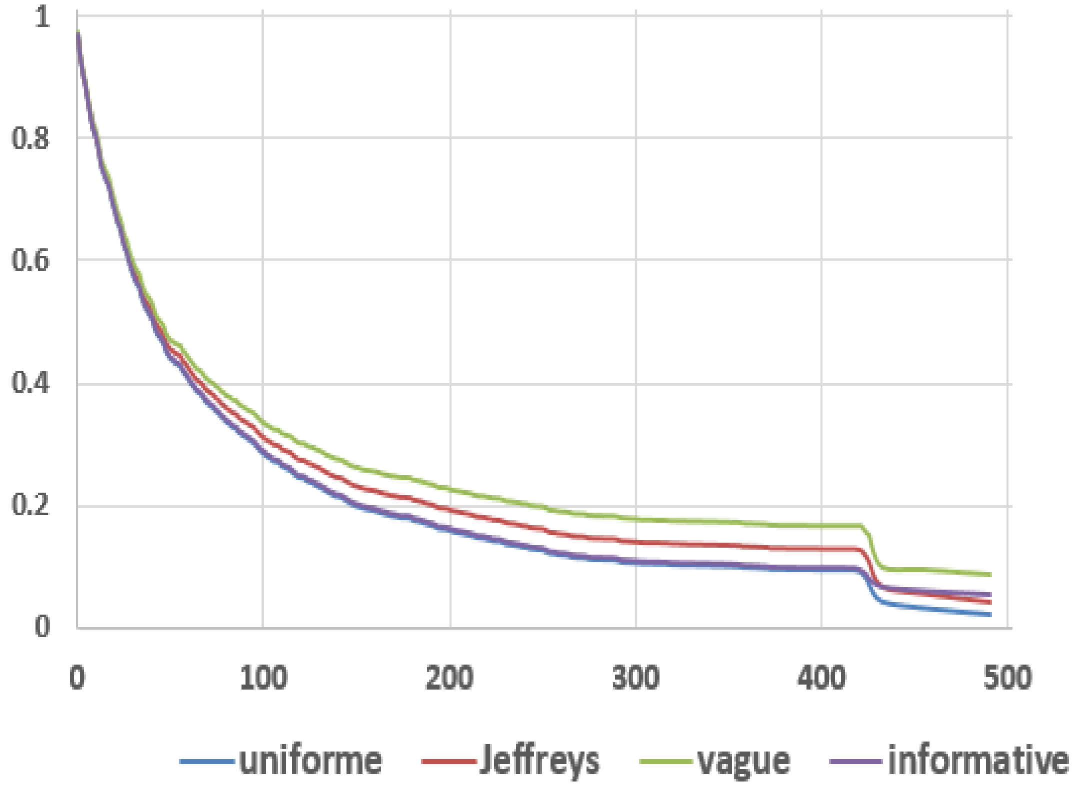
Tab 2 :The comparison between the different Bayesian methods.

the estimation method	**informative**	**uniform**	**Jeffrey's**	**vague**
DIC	1060	1501	1280	1157

Source: Developed by us, using OpenBUGS program

.

Fig 3: The evolution of survival probabilities in the Bayesian Kaplan Meier method according to several a priori distributions.



Source: Developed by us, using OpenBUGS program.

In figure (3) we notice the difference between the estimates of the durations of unemployment where the Bayesian curves represent smoother forms, this figure informs us that the difference occurs gradually after 3 months of unemployment.

VII. Conclusion

In many random experiential situations, the practitioner has information about the phenomenon being studied, opinions of researchers and experts, professional experiences, and observations gained, etc. However, the classical approach somehow ignores this a priori information and only uses observations, in this concept the unknown states are generally considered as quantities with a certain deterministic character. Of course, this drawback is generally erased by asymptotic considerations and a certain number of theorems make it possible to evaluate the good quality of the estimators if the number of observations is sufficient. In the Bayesian measurement approach, an observation transforms this information a priori into a posteriori.

The objective sought in our contribution is to improve the inferential phase in the estimation of nonparametric survival times and in the presence of censorship. The effectiveness of this approach has been demonstrated in a real example by the use of an informative prior distribution.

References

1) Boreux, J., Parent, E., Bernier, J (2010), Pratique du calcule Bayésien. Springer.
2) Colletaz, G (2012), Modèles de survie. Notes de cours, Université d'Orléans, France.
3) Geman, S., Geman, D (1984), Stochastic relaxation, Gibbs distribution and the Bayesian restoration of image. IEEE Trans. Pattern Anal. Mach. Intell., 6,721-741.
4) Horny, G (2007), Hétérogénéité non observée dans les modèles de durée. Discussion Paper, Département des sciences économiques de l'université catholique de Louvain, 46.
5) Robert, C.P. Des spécificités de l'approche bayésienne et de ses justifications en statistique inférentielle. In Les approches et méthodes bayésiennes, sciences et épistémologie (ed. I. Drouet). Éditions Matériologiques (to appear). Available as arxiv:1403.4429.
6) Robert, C.P (2006), Le choix Bayésien : principes et pratiques. Springer.
7) Robert, C.P., Casella, G (2004), Monte Carlo methods. Springer, NewYork, seconde edition.
8) Spiegelhalter, D. J., Best, N., Carlin, B.P., Van der Linde, A (2002), Bayesian measures of model complexity and fit. Journal of the Royal Statistical Society, Series B, 64, 583–640.

Appendices (the code on the OpenBUGS program)

```
model
{
for (i in 1:m)
{
d[i]~dbin(q[i],n[i])
q[i]~dbeta(1,7.77)
}
for (i in 1:m)
{
ce[i]~dbin(0.01,0.01)
}
for (i in 1:m)
{
qc[i]~dbeta(0.01,0.01)
}
```

```
for (i in 1:m)
{
p[i]<-1-q[i]
}
n[1]<- 1064
for(i in 2:m)
{
n[i]<-n[i-1]-d[i-1]-ce[i-1]
}
for (i in 2:m)
{
s[i]<-s[i-1]*p[i]
}
s[1]<-p[1]
}
list( m=227,
ce=c(0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,
0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,
0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,
0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,36,39,29,42,30,6,5,0,1,1),
d=c(28,31,23,18,14,20,18,18,16,8,14,15,20,12,7,9,5,11,15,12,11,13,7,10,16,11,6,13,11,9,9,7,5,4,14,9,9,4,5,6,7,1
2,7,5,3,5,5,11,5,5,3,1,3,1,2,1,8,2,5,4,5,4,4,5,3,1,3,3,6,3,3,4,3,4,1,4,4,1,4,2,1,3,2,1,1,5,4,4,2,2,1,2,3,4,3,4,2,3,2,2,1
,3,1,1,2,4,1,1,1,1,1,3,3,2,3,1,2,2,1,2,2,1,1,2,4,2,1,2,1,3,1,2,3,1,2,2,1,1,1,2,2,1,2,1,1,3,1,1,2,1,1,2,1,2,2,1,2,1,2,1,2,
1,3,1,1,1,2,1,2,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,2,1,1,1,2,2,1,1,1,1,1,1,1,1,1,1,1,1,2,1,1,1,1,1,1,1,1,1,1,0,0,0,0,
0,1,0,1,0,0))
list(q=c(0.5,0.5,0.5,0.5,0.5,0.5,0.5,0.5,0.5,0.5,0.5,0.5,0.5,0.5,0.5,0.5,0.5,0.5,0.5,0.5,0.5,0.5,0.5,0.5,0.5,0.5,0.5,0.5
,0.5,0.5,0.5,0.5,0.5,0.5,0.5,0.5,0.5,0.5,0.5,0.5,0.5,0.5,0.5,0.5,0.5,0.5,0.5,0.5,0.5,0.5,0.5,0.5,0.5,0.5,0.5,0.5,0.5,0.5,0.5,0.5,
0.5,0.5,0.5,0.5,0.5,0.5,0.5,0.5,0.5,0.5,0.5,0.5,0.5,0.5,0.5,0.5,0.5,0.5,0.5,0.5,0.5,0.5,0.5,0.5,0.5,0.5,0.5,0.5,0.5,0.5,
0.5,0.5,0.5,0.5,0.5,0.5,0.5,0.5,0.5,0.5,0.5,0.5,0.5,0.5,0.5,0.5,0.5,0.5,0.5,0.5,0.5,0.5,0.5,0.5,0.5,0.5,0.5,0.5,0.5,0.5,
0.5,0.5,0.5,0.5,0.5,0.5,0.5,0.5,0.5,0.5,0.5,0.5,0.5,0.5,0.5,0.5,0.5,0.5,0.5,0.5,0.5,0.5,0.5,0.5,0.5,0.5,0.5,0.5,0.5,0.5,
0.5,0.5,0.5,0.5,0.5,0.5,0.5,0.5,0.5,0.5,0.5,0.5,0.5,0.5,0.5,0.5,0.5,0.5,0.5,0.5,0.5,0.5,0.5,0.5,0.5,0.5,0.5,0.5,0.5,0.5,
0.5,0.5,0.5,0.5,0.5,0.5,0.5,0.5,0.5,0.5,0.5,0.5,0.5,0.5,0.5,0.5,0.5,0.5,0.5,0.5,0.5,0.5,0.5,0.5,0.5,0.5,0.5,0.5,0.5,0.5,
0.5,0.5,0.5,0.5,0.5,0.5,0.5,0.5,0.5,0.5,0.5,0.5,0.5,0.5,0.5,0.5,0.5,0.5,0.5))
list(q=c( 0.9,0.9,0.9,0.9,0.9,0.9,0.9,0.9,0.9,0.9,0.9,0.9,0.9,0.9,0.9,0.9,
0.9,0.9,0.9,0.9,0.9,0.9,0.9,0.9,0.9,0.9,0.9,0.9,0.9,0.9,0.9,0.9,
0.9,0.9,0.9,0.9,0.9,0.9,0.9,0.9,0.9,0.9,0.9,0.9,0.9,0.9,0.9,0.9,
0.9,0.9,0.9,0.9,0.9,0.9,0.9,0.9,0.9,0.9,0.9,0.9,0.9,0.9,0.9,0.9,
0.9,0.9,0.9,0.9,0.9,0.9,0.9,0.9,0.9,0.9,0.9,0.9,0.9,0.9,0.9,0.9,
0.9,0.9,0.9,0.9,0.9,0.9,0.9,0.9,0.9,0.9,0.9,0.9,0.9,0.9,0.9,0.9,
0.9,0.9,0.9,0.9,0.9,0.9,0.9,0.9,0.9,0.9,0.9,0.9,0.9,0.9,0.9,0.9,
0.9,0.9,0.9,0.9,0.9,0.9,0.9,0.9,0.9,0.9,0.9,0.9,0.9,0.9,0.9,0.9,
0.9,0.9,0.9,0.9,0.9,0.9,0.9,0.9,0.9,0.9,0.9,0.9,0.9,0.9,0.9,0.9,
0.9,0.9,0.9,0.9,0.9,0.9,0.9,0.9,0.9,0.9,0.9,0.9,0.9,0.9,0.9,0.9,
0.9,0.9,0.9,0.9,0.9,0.9,0.9,0.9,0.9,0.9,0.9,0.9,0.9,0.9,0.9,0.9,
0.9,0.9,0.9,0.9,0.9,0.9,0.9,0.9,0.9,0.9,0.9,0.9,0.9,0.9,0.9,0.9,
0.9,0.9,0.9,0.9,0.9,0.9,0.9,0.9,0.9,0.9,0.9,0.9,0.9,0.9,0.9,0.9,
0.9,0.9,0.9,0.9,0.9,0.9,0.9,0.9,0.9,0.9,0.9,0.9,0.9,0.9,0.9,0.9,0.9,0.9,0.9)
```