Finance and Business Economics Review | Volume 4 / Number 4 / December 2020/ pp150-168

The Bayesian Design of Kaplan Meier Estimation Using Gibbs Sampling: Application in econometrics of duration models

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

[1] Phd Student, National Higher School of statistics and applied economics, Tipaza, Algeria,ahmedhamimes@yahoo.com.

[2] Professor, National Higher School of statistics and applied economics, Tipaza, Algeria,,rbena2002@hotmail.com

Received: 28/11/2020 **Accepted**: 18/12/2020 **Published**: 31/12/2020

Abstract: A Bayesian approach to survival offers practical, simple and relatively easy solutions to exploit digitally. In this contribution, we will demonstrate the effectiveness of the Bayesian approach in the modeling of durations and in an econometric context, we propose the Bayesian design of the Kaplan Meier estimator based on the stochastic approximation, which is represented here by the Gibbs sampling. Our contribution is to improve the deductive stage in estimating nonparametric survival times and under censorship, and this is what we reached in our research by means of the hierarchical prior distribution.

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

(JEL) Classification : C11, C15, C41, E24.

Résumé : Une approche bayésienne de la survie offre des solutions pratiques, simples et relativement faciles à exploiter numériquement. Dans cette contribution, nous démontrerons l'efficacité de l'approche bayésienne dans la modélisation des durées et dans un contexte économétrique, nous proposons le plan bayésien de l'estimateur de Kaplan Meier basé sur l'approximation stochastique, qui est ici représentée par l'échantillonnage de Gibbs. Notre contribution est d'améliorer l'étape déductive dans l'estimation des temps de survie non paramétriques et sous censure, et c'est ce que nous avons atteint dans notre recherche au moyen de la distribution hiérarchique a priori.

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

(JEL) Classification: C11, C15, C41, E24.

[*] Corresponding author.

1. INTRODUCTION

The analysis of censored lifetimes is used in various fields of application and different possibilities have been proposed for the modeling of such data. The first field of application of the analysis of duration data was in the biomedical sciences where it is used either for a therapeutic trial or for epidemiological studies. In economics, we study the length of time spent unemployed, in a job or between two jobs, the duration of a transport trip, the lifespan of a business or the duration of a "revolving" type loan.

In Bayesian inference, survival analysis has received increasing attention in recent years, but still limited due to the scarcity of specialized software (one of the causes of this scarcity is the difficulty in automating Bayesian analyzes compared to the frequentist approach), and also the force of habit and the difficulty in accepting a different conception of statistics. A great interest of Bayesian inference over frequentist methods is its great consistency and its unified methodology in the theoretical plan, which allows to deduce results of richer and more direct interpretations than those provided by the classical approach.

In classical nonparametric modeling of survival times in most of the situations encountered today in clinical research, the Kaplan-Meier method seems more appropriate and more precise. In some special cases, the actuarial method can still retain indications for use today in the event that the times of occurrence of unknown events or the study size are high. In this contribution, we will give a Bayesian alternative of the actuarial estimator and classical Kaplan Meier based on the beta distribution, after we present several methods to calculate the prior distribution. We find this important to give a Bayesian estimator based on these two classical conceptions for several reasons, firstly and in medical science for example, we very often have to deal with small samples. There are several reasons for this. The most common is the rarity of the disease or the difficulty in bringing together patients with the same biochemical parameters. In addition, we have very often censored data. Therefore, a small sample size does not allow us to use classical statistical methods or when they are used they can give us too general and even false results. The second reason by using the Bayesian method, we get results for both censored and uncensored cases. This is why our survival curve is smoother and we don't have such rapid jumps for the probability of survival. In addition, for some observations, the interval is even smaller. The resort to the Bayesian method in the case of the Kaplan Meier or Actuarial estimator simplifies the use of this approach compared to other methods (methods based on randomized measurement) and addresses the problem of strength of habit. in the use of frequentist estimates. Several works have been based on the development of the Kaplan Meier estimator. Rossa and Zieliński (1999) used a solution based on an approximation by the Weibull distribution function as a local smoothing of the Kaplan-Meier estimator. Rossa and Zieliński (2002), uses Weibull's law as an approximation function of the Kaplan Meier method. Shafiq Mohammad et al (2007), presented a weighting of the Kaplan Meier estimator for heavily censored data under the sine function. Khizanov and Maĭboroda (2015), used a mixing model with varying concentrations as a change or modification of the Kaplan Meier estimator.

In our application we will analyze the durations of global unemployment in the National Employment Agency (ANEM) of Ain El Benian. We are working on a sample of 1064 unemployed individuals observed between 01/01/2011 and 07/15/2013. This application allows to demonstrate practically that the Bayesian procedures in econometrics constitute an essential element for the control of the economic information because of the difference which exists in the interpretation and the estimate of the curves and the durations of exit from unemployment, this difference gives a form of a set of options with differential information, on the other hand the frequentist approach which is presented in this example as a particular case of Bayesian inference.

2. BAYESIAN STATISTICAL MODEL

The concept of “Bayesian statistics” starts from the neologism “Bayesian”, taken from the name of Thomas Bayes, who introduced the theorem which now bears his name in a posthumous article of 1763, 250 years ago. This theorem expresses a conditional probability in terms of the inverse conditional probabilities, if A and E are events such that $P(E) \neq 0$, $P(A/E)$ and $P(E/A)$ are connected by:

$$P(A/E) = \frac{P(E/A)P(A)}{P(E)},$$

Bayes used the continuous version of this theorem, taking two random variables x and y. The conditional distribution of y given x is given by:

$$h(y/x) = \frac{f(x/y) \times f(y)}{\int f(x/y) \times f(y) dy} \qquad (1)$$

This result makes it possible to carry out the inference from the law of the parameter θ conditional on the observations x, called a posteriori law and defined by:

$$\pi(\theta/x) = \frac{f(x/\theta) \times \pi(\theta)}{\int_{\Theta} f(x/\theta) \times \pi(\theta) d\theta} = \frac{f(x/\theta) \times \pi(\theta)}{m(x)} \qquad (2)$$

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:

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

3. Monte Carlo Marko chain method

From a point of view and when the explicit calculation of the posterior distribution of the vector of the parameters to be estimated is particularly difficult by hand, the Monte Carlo methods by Markov Chains make it possible to make simulations 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)).

4. BASIC CONCEPTS IN SURVIVAL ANALYSIS

Let X be a positive or zero random variable, and absolutely continuous defined on a probability space (Ω,A,P) which describes the time that elapses between two events.

The hazard rate (or risk function, failure rate, failure rate, death rate, instantaneous risk, etc.) at time t is called the function:

$$\lambda(t)=\lim_{dt\to 0}\frac{1}{dt}P(t\le X<t+dt/X\ge t),t>0$$

In continuous time, the risk function is the limiting probability that a unit will make a transition knowing that it has not made one beforehand. We can show the following relation:

$$\lambda(t)=\frac{f(t)}{S(t)}$$

When t fixed, characterize the probability of dying in a small time interval after time t, conditional on having survived until time t, so this means the risk of instantaneous death for those who survived.

4.1. KAPLAN MEIER'S METHOD

4.1.1. Definition of Kaplan-Meier method

The Kaplan-Meier (KM) estimation method is also called by Anglo-Saxon statisticians "Product Limit Estimations (PLE)". This estimator, which is a generalization of the notion of empirical distribution function, is based on the following idea: to survive after time t is to be alive just before t and not to die at time t.

If X, a random variable, represents the time elapsed since an instant t_0 and when the time is considered in a discrete manner, if t_i represents an instant during which there is the observation of at least one event, then the probability of survival at time t_i is equal to the probability of having survived before t_i multiplied by the “conditional” probability of surviving at time t_i. The use of the term "conditional" here means that it is the probability of surviving time t_i knowing that the individuals were survivors in t_i:

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

account. Let us call d_i and c_i, the numbers of individuals who, respectively, know the event and exit from observation at t_i. The number n_i of individuals subject to the risk of knowing the event at t_i corresponds to the set of individuals who, just before this instant t_i was reached, had neither known the observed event, nor n 'had come out of observation.

Furthermore, in the case of the Kaplan-Meier estimate, we consider that the exits from observation c_i take place a fraction of the time after the deadlines d_i (Le Goff et al, 2013). Therefore, 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 $t = t_i$.

The probability of survival at t_i then becomes:

$$S(t_i) = S(t_{i-1})(1 - q_i)$$

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

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 observation:

$$\hat{S}(t) = \begin{cases} \prod_{t_i \leq t} \left(1 - \frac{d_i}{n_i}\right) = \prod_{t_i \leq t} (1 - q_i) & \text{si } t \geq t_1 \\ 1 & \text{si } t < t_1 \end{cases} \quad (4)$$

4.1.2. The Bayesian conception of Kaplan Meier estimator

A. The beta-Binomial model

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

$$q_i = 1 - \frac{d_i}{n_i} \quad (5)$$

From a Bayesian perspective we assume an a priori for q_i, and when the distribution used in the case of proportions is that of Beta, we set:

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

For the hyperparameters (α, β), we find several propositions:

B. 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:

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

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 $\pi\left(q_i/d_i,\hat{\alpha},\hat{\beta}\right)$.

also let :

$$\begin{cases} n_1 = \text{the number of subjects at the start of the study} \\ n_i = n_{i-1} - d_i - c_i \end{cases},$$

for the application see code A1.

C. **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

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

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

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 :

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

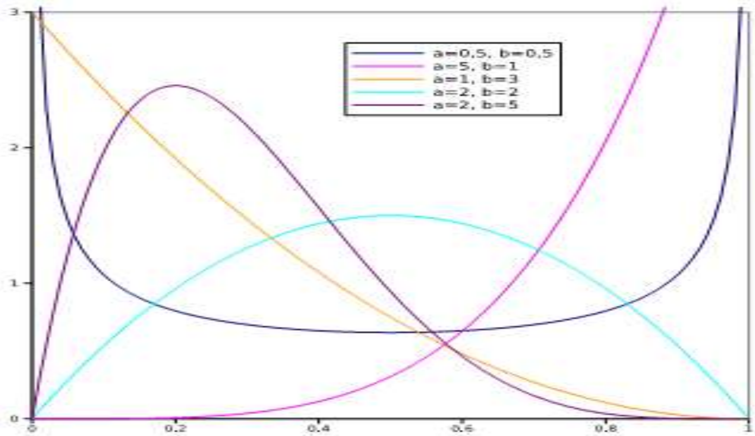
starting from A, the a priori function is written as follows:

$$\pi(q_i) = A * \pi^*(q_i) = \frac{1}{\sqrt{n}\beta(1/2;1/2)} n^{1/2}\left(q_i(1-q_i)\right)^{1/2-1}$$
$$= \frac{1}{\beta(1/2;1/2)} q_i^{-1/2}(1-q_i)^{-1/2}; 0 < q_i < 1$$

so

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

Table (01): the distribution of beta (a ; b)



Source: Developed by us.

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.

also let :

$$\begin{cases} n_1 = \text{the number of subjects at the start of the study} \\ n_i = n_{i-1} - d_i - c_i \end{cases},$$

for the application see code A2.

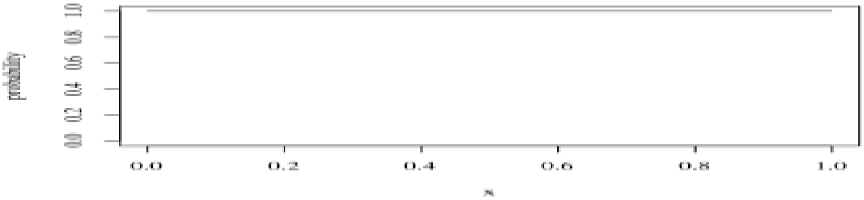
D. **Under the a priori uniform law**

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:

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

Figure (2): the distribution of beta (1; 1)



Source: Developed by us.

also let :

$$\begin{cases} n_1 = \text{the number of subjects at the start of the study} \\ n_i = n_{i-1} - d_i - c_i \end{cases},$$

for the application see code A3.

E. The hierarchical link function

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:

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

and

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

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

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. τ follows a half-Cauchy law if that density is:

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

on note

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

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

The proposed model is written as follows

$$q_i = \frac{exp(\mu_i)}{1 + exp(\mu_i)}$$
$$\mu_i \sim \mathcal{N}(\vartheta;\tau)$$
$$\vartheta \sim \mathcal{N}(0;0{,}001),\ \tau \sim HC(B)$$
$$B \sim Uniforme(0;\mathcal{T}),\text{we pose}: \mathcal{T} = 100.$$

also let :

$$\begin{cases} n_1 = \text{the number of subjects at the start of the study} \\ n_i = n_{i-1} - d_i - c_i \end{cases},$$

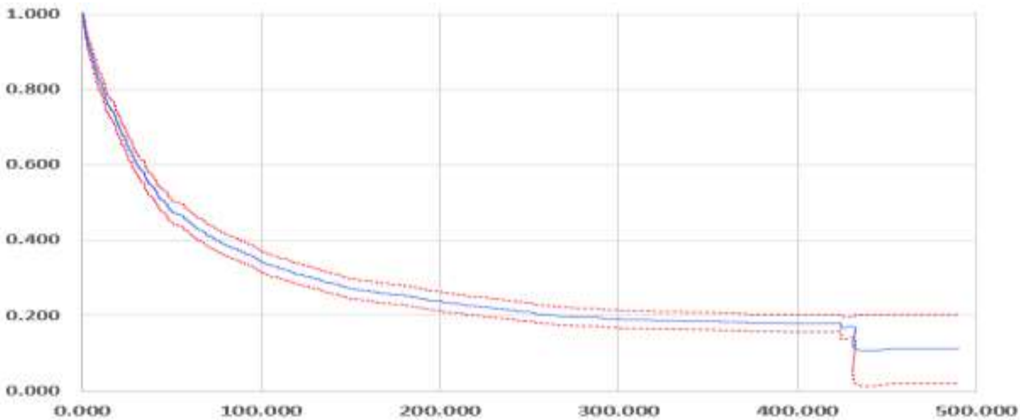
for the application see code A4.

5. Application in econometrics of duration models

- **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 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.

Figure (3): Kaplan-Meier survival functions for overall unemployment duration.

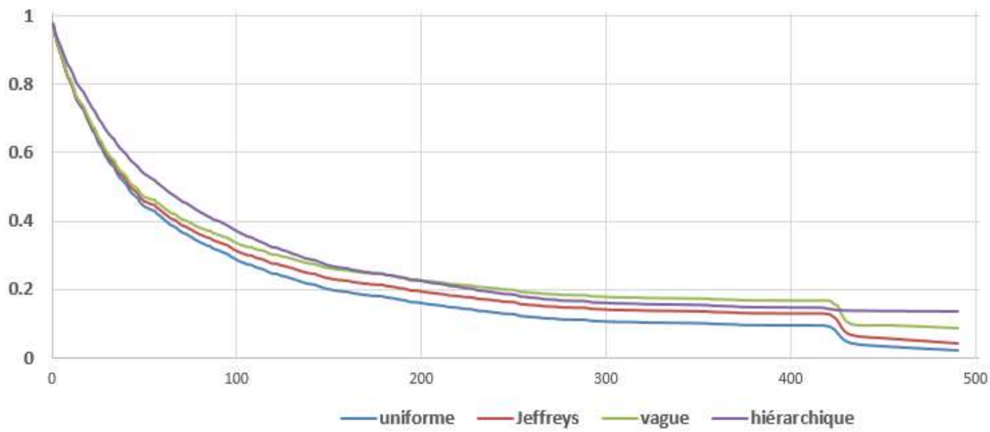


Source: Developed by us.

From Figure (3), 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 (4) also informs us that the median duration of exit from unemployment is 45 days in both methods. Thus, the two classical estimation methods (actuarial and Kaplan-Meier) are identical approximations although the number of individuals is large, and in the aspect of precision we find that the actuarial method has a higher efficiency, in what follows we will compare it with the Bayesian methods.

Figure (4): Bayesian Kalan Meier survival functions according to several a priori distributions.



Source: Developed by us.

In figure (4), we notice a similarity between the survival curves according to the different propositions of the prior law. Consequently, the DIC (the information deviance criterion) is used to choose the best model.

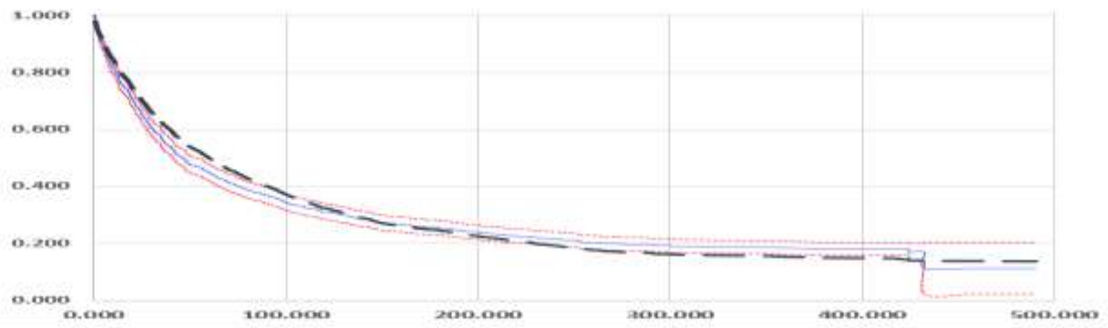
Table (1): The comparison between the different Bayesian methods for the KM.

Kaplan Meier method	Jeffreys	uniform	vague	hierarchical
DIC	1280	1501	1157	920

Source : Developed by us.

In table (1) the Kaplan Meier model with a Hierarchical prior law is preferable compared to the other laws used in this comparison.

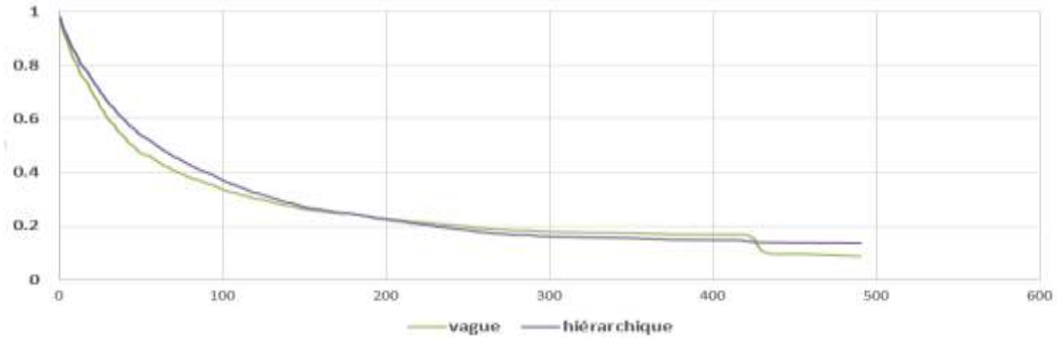
Figure (5): The evolution of survival probabilities in the classical method (in blue) of KM and Bayesian with a vague prior law (the discontinuous curve).



Source: Developed by us.

In Figure (5) we notice a small difference between the estimates of unemployment durations between the Bayesian and classical method, even at the median level (45 days). Bayesian curves represent smoother shapes compared to the frequentist or classical method. Consequently, the frequentist approach presents in this example (the Kaplan Meier model and the Bayesian model with a vague a priori of beta (0.01, 0.01)) a particular case of Bayesian inference.

Figure (6): The evolution of survival probabilities in the vague and hierarchical Bayesian KM method.



Source: Developed by us.

In figure (6) we notice a small difference between the estimates of unemployment durations between the vague and hierarchical Bayesian method, with a level of median (45 days) and (60 days) respectively, the latter can generate a multitude of decisions between approaches.

6. Results

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. In the results of this work, we find:

The relevance and effectiveness of the Bayesian approach as a guide to scientific reasoning in the face of uncertainty has long been recognized, we have shown this effectiveness with the application of Gibbs sampling. The use of the MCMC method is relatively easy to implement, it provides a set of techniques very suitable for the estimation of complex models with several parameters or with a hierarchical structure. The use of the MCMC method is relatively easy to implement, it provides a set of techniques very suitable for the estimation of complex models with several parameters or with a hierarchical structure.

- From the results of the application we find that the choice between the two statistical approaches (Bayesian and classical) is likely to generate a multitude of decisions.

- The Bayesian approach and through the a priori distribution is richer compared to the frequentist method.
- The frequentist approach presents in this example (the Kaplan Meier model and the Bayesian model with a vague a priori of beta (0.01, 0.01)) a particular case of Bayesian inference.

7. 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, unknown states are generally considered to be 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.

8. References

- Geman, S., Geman, D. (1984). Stochastic relaxation, Gibbs distribution and the Bayesian restoration of image. IEEE Trans. Pattern Anal. Mach. Intell., 6,721-741.
- Khizanov, V. G., Maĭboroda, R. (2015). A modified Kaplan-Meier estimator for a model of mixtures with varying concentrations. Theor. Probability and Math. Statist. 92 (2016), 109-116.
- Le Goff, J.M., Foney, Y. (2013). Méthodes non paramétriques de l’analyse des événements du parcours de vie (event history analysis). Cahier de Recherches et Méthodes, Université de Lausanne, 2.
- 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.
- Robert, C.P. (2006). Le choix Bayésien : principes et pratiques. Springer.

- Robert, C.P., Casella, G. (2004). Monte Carlo methods. Springer, New York, seconde edition.
- Rossa, A. and Zieli´nski, R. (1999). Locally Weibull-Smoothed Kaplan–Meier Estimator, Institute of Mathematics Polish Academy of Sciences, Preprint 599.
- Rossa, A., & Zieliński, R. (2002). A simple improvement of the Kaplan-Meier estimator. Communications in Statistics-Theory and Methods, 31(1), 147-158.
- Shafiq, M., Shah, S., & Alamgir, M. (2007). Modified Weighted Kaplan Meier Estimator. Pakistan Journal of Statistics and Operation Research, 3(1), 39-44.