Journal of Economic Growth and Entrepreneurship JEGE
Spatial and entrepreneurial development studies laboratory
Year : 2020 Vol.5 No.1 pp:1-10.

Bayesian methods in applied econometrics, or advantages and properties of Bayesian statistics

Hamimes Ahmed [*1], *Benamirouche Rachid* [2], *Chellai Fatih* [3]
[1] *MAA, statistics Department, National School of Statistics and Applied Economics, Algeria.*
[2] *Professor, statistics Department, National School of Statistics and Applied Economics, Algeria.*
[3] *MCB, basic education, Faculty of Economics, Business and Management, University of Ferhat Abbas, Algeria*

ARTICLE INFO
Article history:
Received:30/08/2020
Accepted:9/11/2020
Online:27/12/2020

Keywords:
Bayesian statistics.
Bayesian econometrics
Characteristics.
Advantages.
properties.
JEL Code:C11 ;C02.

ABSTRACT

Bayesian statistics can be presented as a generalization of the frequentist approach. The choice of probabilized the space of states is truly revolutionary. Statisticians draw a hermetic border between the notion of unknown parameter and random parameter. In this article, we want to show the interest of Bayesian econometrics and this analytical approach on three important points; in the level of characteristics, in the advantages and finally in the properties.

Méthodes bayésiennes en économétrie appliquée, ou avantages et propriétés de la statistique bayésienne

Hamimes Ahmed [1], *Benamirouche Rachid* [2], *Chellai Fatih* [3]
[1] *École nationale de la statistique et d'économie appliquée, Algérie.*
[2] *École nationale de la statistique et d'économie appliquée, Algérie.*
[3] *Faculté d'économie, de commerce et de gestion, Université de Ferhat Abbas, Algérie*

ARTICLE INFO
Reçu: 30/08/2020
Accepté: 9/11/2020
En ligne: 27/12/2020

Mots clés :
statistique Bayésienne.
l'économétrie bayésienne.
caractéristique.
avantages.
propriétés.
Code JEL: C11 ;C02.

RÉSUMÉ

la statistique Bayésienne peut être présentée comme une généralisation de l'approche fréquentiste. Le choix de probabilisé l'espace des états est vraiment révolutionnaire. Les statisticiens tracent une frontière hermétique entre la notion de paramètre inconnu et de paramètre aléatoire. Dans cet article, nous souhaitons montrer l'intérêt de de l'économétrie bayésienne et cette approche d'analyse sur trois points importants ; dans le niveau des caractéristiques, dans les avantages et finalement dans les propriétés.

[*] *Corresponding Author e-mail : ahmedhamimes@yahoo.com.*

Introduction

Thomas Bayes[1] (1702-1761) est un grand mathématicien qui n'est pas passé à l'histoire pour la découverte de la formule qui porte son nom, mais pour son interprétation et son application à un problème d'estimation. Il a inspiré depuis des siècles les statisticiens, les mathématiciens, les physiciens et même d'autres chercheurs aux sciences sociales, ce qui donne aujourd'hui une approche cohérente et pratique pour résoudre les problèmes d'inférence appelée l'approche Bayésienne. Le fondement du paradigme Bayésien s'est basé sur la théorie d'inversion des probabilités connue sous le nom du théorème de Bayes. À partir de cette formule, il est possible de formaliser l'inversion des conditionnements dans les probabilités. Laplace et Bayes sont allés plus loin et ont considéré que le paramètre peut prendre plusieurs valeurs possibles, avec des probabilités associées, ce qui conduit à trouver des distributions sur l'espace des paramètres, où chaque distribution de probabilité est appelée distribution a priori. L'analyse statistique Bayésienne consiste à déduire de cette loi de probabilité, grâce au théorème d'inversion, une distribution dite « a posteriori ». En 1774, Pierre Simon de Laplace approfondit la notion d'inversion en probabilité à partir d'un changement de la justification physique des distributions a priori introduites par Bayes, à travers des justifications fondées sur un raisonnement abstrait, de1931 et 1937, de Finetti suivit les travaux de la communauté Bayésienne et introduit la notion d'échangeabilité et en 1972 a démontré qu'il existe une mesure de probabilité supporte leur théorème fondamental et justifié l'utilisation des méthodes Bayésiennes. Tout ce passage va donner un statisticien qui raisonne différemment, puisqu'il va probabiliser tout ce qui est incertain et même des phénomènes non aléatoires, il va chercher à quantifier son incertitude à travers toutes les informations disponibles. Cette distinction nous permet de comprendre pourquoi le paradigme Bayésien a également sa place en Biostatistique , économie ,…. Le concept bayésien se différencie du concept classique dont le sens où le paramètre est une variable aléatoire dont le comportement est supposé connu, en lui associant une loi de probabilité sur l'espace Θ dite loi a priori et notée $\pi(\theta)$, et à travers cette conception l'analyse statistique permet de considérer toutes les informations qualitatives et quantitatives sur l'incertitude dans le modèle. Ensuite, si on utilise le théorème de Bayes qui permet d'inverser les probabilités, on peut déduire la distribution a postériori $\pi(\theta/x)$ qui nous permet de construire des procédures inférentielle de la manière la plus naturelle possible, ce qui explique aussi la persistance de ce paradigme, contre vents et marées depuis 250 ans. De nombreux ouvrages ont été publiés sur les statistiques bayésiennes, notamment Bernardo et Smith (1994), O'Hagan (1994), Carlin et Louis (2001), Gelman et al. (2003), O'Hagan et Forster (2002) et Schervish, (1995).

Dans cet article, nous présentons les fondements de base de cette approche et nous souhaitons montrer l'intérêt de ce paradigme d'analyse en économie sur trois points importants ; dans le niveau des caractéristiques, dans les avantages et finalement dans les propriétés. Dans les références de l'économétrie bayésienne, on trouve Koop, G. (2003), Koop et al (2007) et Geweke (2005).

[1] Thomas Bayes (1702 ?-1761) était un prêtre presbytérien non-conformiste, membre de la Roy al Society, qui publia à titre posthume « An Essay to wards solving a Problem in the Doctrine of Chances ». On connaît très peu de détails sur sa vie et le seul portrait de lui dont on dispose demeure incertain.

I. L'économétrie bayésienne

1. Définition

L'économétrie bayésienne est une branche de l'économétrie qui applique la modélisation économique aux principes bayésiens. Le bayésianisme est basé sur une compréhension de la probabilité du degré de conviction, par opposition à une compréhension de la fréquence relative.

2. Thèmes de recherche actuels

Depuis le début du 21e siècle, la recherche en économétrie bayésienne s'est concentrée sur (Basturk, N, (2013)) :

- méthodes d'échantillonnage adaptées à la parallélisation et aux calculs GPU;
- des modèles économiques complexes tenant compte des effets non linéaires et des densités prédictives complètes;
- analyse des caractéristiques implicites du modèle et analyse des décisions;
- intégration de l'incomplétude du modèle dans l'analyse économétrique.

II. FORMULE DE BAYES ET LOI A POSTERIORI

Dans le cas discret et si on considère $(H_i)_{i\in(1\dots n)}$ l'ensemble de toutes les hypothèses (ou bien les causes ou les circonstances) de réalisation d'un événement E de probabilité non nulle. Si l'effet E se réalise en même temps qu'un et un seul des H_i , c.à.d.

$$E = (E\cap H_1)\cup(E\cap H_2)\cup \dots \cup(E\cap H_n),$$

selon la théorie des probabilités totales on écrit

$$P(E) = P(E\cap H_1) + P(E\cap H_2) + \cdots + P(E\cap H_n)$$

$$= \sum_{i=1}^{n} P(E\cap H_1) = \sum_{i=1}^{n} P(E/H_i)P(H_i) = E^H\big(P(E/H_i)\big),$$

la probabilité conditionnelle d'une cause H_i est donnée par

$$P(H_i/E) = \frac{P(E\cap H_i)}{P(E)} = \frac{P(E/H_i)P(H_i)}{E^H\big(P(E/H_i)\big)}, \qquad (1)$$

l'équation (1) est la théorie (ou la règle) de Bayes.

Nous supposons une séquence additionnelle E' et d'après la théorie de Bayes, on écrit :

$$P(H_i/E',E\ ) = \frac{P(E',E\ /H_i)P(H_i)}{E^H\big(P(E',E\ /H_i)\big)}$$

$$\propto P(E',E\ /H_i)P(H_i)$$

$$\propto P(E'\ /E,H_i)P(E\ /H_i)P(H_i)$$

$$\propto P(H_i\ \ /E)P(E'\ /E,H_i),$$

d'après cette règle nous constatons que ce théorème est un principe d'actualisation, il décrit comment la connaissance modifie l'expérience. Si on suppose que E et E' sont conditionnellement indépendantes, donc

$$P(E'\ /E,H_i)\ P(E\ \ /H_i) = P(E'\ /H_i)\ P(E\ /H_i)$$

Si nous avons m expériences, sous l'hypothèse de l'indépendance conditionnelle on trouve

$$P(H_i/E^m, E^{m-1}, \dots, E^1) \propto \prod_{j=1}^{m} P(E^j/H_i)\, P(H_i)$$

Pour E l'expérience complète on écrit

$$P(H_i/E) \propto exp\left\{\sum_{j=1}^{m} log[P(E^j/H)] + logP(H_i)\right\},$$

on pose

$$L = \frac{1}{m}\sum_{j=1}^{m} log[P(E^j/H_i)],$$

donc

$$\begin{aligned} P(H_i/E) &\propto exp\{mL + logP(H_i)\} \\ &\propto exp\left\{mL\left[1 + \frac{logP(H_i)}{mL}\right]\right\}, \end{aligned}$$

pour une grande valeur de m on trouve que :

$$\begin{aligned} P(H_i/E) &\propto exp(mL) \\ &\propto exp\left(\sum_{j=1}^{m} log[P(E^j/H_i)]\right) \\ &\propto \prod_{j=1}^{m} P(E^j/H_i), \qquad (2) \end{aligned}$$

l'équation (2) indique que la contribution de la probabilité a priori au probabilité a postériori est d'ordre $1/m$. D'une autre manière, plus les informations des épreuves s'accumulent, plus le rôle de l'a priori dans la probabilité a postériori se diminué.

Le concept de « statistique bayésienne » part du néologisme « bayésien », tiré du nom de Thomas Bayes (utilisé pour la première fois dans les années 1950), qui introduisit le théorème qui porte à présent son nom dans un article posthume de 1763, il y a 250 ans. Bayes utilisait la version continue de ce théorème, en prenant deux variables aléatoires x et y. La distribution conditionnelle de y sachant x est donnée par :

$$h(y/x) = \frac{f(x/y) \times f(y)}{\int f(x/y) \times f(y)dy}, \qquad (3)$$

l'expression (3) permet de mener l'inférence à partir de la loi du paramètre θ conditionnellement aux observations$x$, appelée loi a posteriori et définie par :

$$\pi(\theta/x) = \frac{f(x/\theta) \times \pi(\theta)}{\int_{\Theta} f(x/\theta) \times \pi(\theta)d\theta} = \frac{f(x/\theta) \times \pi(\theta)}{m(x)} \qquad (4)$$

cette loi a postériori est la combinaison de

- $f(x/\theta)$ la fonction de densité de x sachant la valeur de la variable aléatoire θ.
- $\pi(\theta)$ modélise la fonction de densité a priori sur θ.
- $m(x)$ la distribution marginale de x.

L'expression (4) représente ce qu'on sait et ce qu'on ne sait pas avant par rapport au paramètre en considérant les données observées; c'est aussi la mise à jour de $\pi(\theta)$ après l'observation de notre échantillon.

Une fois que l'on dispose des données, la quantité $m(x)$ est une constante de normalisation qui garantit que $\pi(\theta/x)$ est bien une distribution de probabilité. On peut écrire :

$$\pi(\theta/x) \propto f(x/\theta) \times \pi(\theta) \qquad (5)$$

l'équation (5) montre que l'inférence bayésienne vérifie le principe de vraisemblance : a posteriori, l'information provenant des données provient exclusivement de la vraisemblance $f(x/\theta)$.

1. Le modèle bayésien

Soit X un vecteur aléatoire à valeurs dans IR^n défini sur un espace probabilisé (Ω, τ, IP), o'u l'univers Ω est un ensemble quelque que, l'ensemble des événements τ est une tribu sur Ω, et IP set une probabilité sur τ. la loi de X n'est que partiellement connue. On suppose que cette loi appartient à une classe donnée $\{P_\theta, \theta \in \Theta\}$ de loi de probabilité, le but de l'analyse statistique est de faire l'inférence sur θ, i.e. il existe $\theta_0 \in \Theta$ tel que $X \sim P_{\theta_0}$. Posons $\chi = X(\Omega) \subset IR^n, \chi$ est l'espace des observations, P_θ est une loi sur χ muni de $\mathcal{F}$ (sa tribu borélienne).

Définition 1. (Le modèle statistique).

On appelle **modèle statistique** ou structure statistique, le triplet $(\mathfrak{X}, \mathcal{F}, P_\theta, \theta \in \Theta)$. on écrira aussi (Ω, P_θ).

Définition 2. (Le modèle statistique bayésien).

On appelle **modèle statistique bayésien**, la donnée d'un modèle statistique paramétré $(\mathfrak{X}, \mathcal{F}, P_\theta, \theta \in \Theta)$ avec $f(x/\theta)$ densité de P_θ et d'une loi $\pi(\theta)$ sur le paramètre.

- La modélisation bayésienne est paramétrique si la famille de départ est paramétrée par un vecteur de dimension fini.
- La modélisation bayésienne est non paramétrique s'il s'agit d'estimer un vecteur de dimension infinie. D'une manière générale

$$\Theta \in IR^n \rightarrow modèle\ paramétrique$$
$$\Theta\ plus\ gros \rightarrow modèle\ non\ paramétrique$$

2. Pourquoi l'approche bayésienne

- La statistique classique définit la probabilité comme une proportion limite déterminée (par la stabilisation des fréquences) sur la base d'une séquence infinie d'expériences, mais dans plusieurs situations on calcule les probabilités à partir des évènements qui par définition ne sont pas repérables, ou en présence de très grandes bases de données, le paradoxe est que tout devient significatif, car l'inférence statistique classique a été développée pour traiter des « petits » échantillons. En revanche, la probabilité en statistique bayésienne est la quantification d'un degré d'incertitude et permet de combiner plusieurs sources d'information sur l'incertitude dans le modèle.
- L'analyse bayésienne fournit des résultats d'interprétation plus riches (à traves l'explicitation du conditionnement probabiliste) que ceux de la statistique classique.
- L'analyse bayésienne permet l'emploi naturel des concepts liés que sont l'échangéabilité et les structures hiérarchiques (Parent et Bernier, 2007).
- Facile à faire des prédictions à traves l'utilisation directe de la distribution prédictive a priori ou la vraisemblance marginale. Plus que ca, la méthodologie bayésienne n'utilise pas deux fois les données pour un problème de prévision.
- Cette approche nous donne les valeurs plausibles pour le paramètre d'intérêt, ce qui rendre l'interprétation plus profonde et facilité l'élaboration des différentes tests statistiques.

- L'interprétation bayésienne est naturelle, l'exemple le plus flagrant est l'intervalle de crédibilité, définie comme la probabilité que θ est dans l'intervalle C, si on répète l'expérience n fois. D'après Dalta et Mukerjee (2004) et pour les intervalles de crédibilités unilatéraux et les régions HDR, $P_\theta(\theta \in C) = 1 - \alpha + 0(1)$ quand n tend vers ∞. En revanche l'intervalle de confiance fréquentiste définit par :

$$P(I_1 \leq \theta \leq I_2) = 1 - \alpha,$$

la méthode fréquentiste cherche à $(1-\alpha)100\%$ intervalles de confiances aléatoires (d'amplitude aléatoire) recouvrent la valeur d'un paramètre inconnue, et lorsque cet intervalle dépend des données la probabilité de couverture diffère en générale de$(1-\alpha)$.

- Parmi les algorithmes de calcul d'inférence les plus efficaces fondés sur les méthodes MCMC, on trouve l'algorithme de Métropolies-Hastings ou encore l'algorithme de Gibbs, sont apparus il y a 50 ans pour la physique statistique, ont des applications presque illimitées, même si ses performances varient largement, selon la complexité du problème, ils ont à cet effet provoqué des développements spectaculaires récents de la statistique bayésienne, et permettent d'estimer les distributions a postériori de modèles complexes, de grande dimensions ou à structure hiérarchique. En effet l'estimation de ces modèles avec les méthodes classiques se révèle parfois nettement plus compliquée.

III. LES CARACTERISTIQUES DE L'APPROCHE BAYESIENNE

- **L'information a priori**

La conception bayésienne ne remet pas en cause l'existence de plusieurs phénomènes ou le paramètre est naturellement inconnue est non aléatoire, la représentation bayésienne propose une méthode de présentation d'information disponible sur θ, plutôt qu'une conception physique du paramètre aléatoire.

- **Approche unificatrice**

Le modèle bayésienne est caractérisé par une mesure de probabilité unique sur d'espace produit (observations * paramètre). En reverche un seul outil, le théorème de Bayes, qui est utilisé dans toutes les situations. Cela conduits à une réponse unitaire et globale, dans tous les cadres de statistiques (paramétriques, non paramétrique et semi-paramétrique).

- **Le principe d'actualisation (les données changes les probabilités)**

D'après l'équation (1.1), la distribution a posteriori représente l'actualisation de l'information disponible sur le paramètre θ au vu de l'information contenue dans la vraisemblance $f(x/\theta)$, tandis que $\pi(\theta)$ représente la connaissance disponible a priori. Le cadre bayésien se présente donc comme une théorie formalisée de l'apprentissage par l'expérience.

- **Le principe d'inversion**

D'après cette approche et lorsque et lorsque θ , X constituent deux variables aléatoires, $f(X/\theta)$ est une densité de probabilité conditionnelle. Selon cette dernière la vraisemblance dans la statistique bayésienne est une densité de probabilité rendre cette approche la seul axiomatique respecte le principe d'inversion en statistique (elle vise à déterminer les causes a partir des effets).

- **La marginalisation**

La marginalisation permet à l'analyse bayésienne d'éliminer les paramètres de naissance, d'une autre manière elle réduit la dimension de l'espace des paramètres à estimer. Si on suppose $\theta = (\theta_1, \theta_2) \in \Theta$, et θ_2 un paramètre de nuisance, alors la densité marginale de $\pi(\theta_1/x)$ s'écrit par l'intégration de la fonction $\pi(\theta/x)$ sur le domaine de définition du paramètre de nuisance comme suit

$$\pi(\theta_1/x) = \int_{\Theta_2} \pi(\theta/x) d\,\theta_2$$

IV. LES PROPRIETES ASYMPTOTIQUES DE LA DISTRIBUTION A POSTERIORI

Dans cette section, on parle du comportement asymptotique de la distribution a posteriori. Le résultat de convergence peut être utilisé comme une approximation de la loi a postériori si la taille de l'échantillon est suffisamment grand, et comme une validation fréquentiste da la statistique bayésienne.

1. La convergence[2]

a) Le cas discret

Dans le cas de l'espace paramétrique discrète, on écrit :

$$\pi(\theta_i/X) = \frac{f(X/\theta_i) \times \pi(\theta_i)}{\sum_i f(X/\theta_i) \times \pi(\theta_i)},$$

si on pose θ_0 le vrai paramètre, on peut écrire :

$$\pi(\theta_i/X) = \frac{\frac{f(X/\theta_i)}{f(X/\theta_0)} \times \pi(\theta_i)}{\sum_i \frac{f(X/\theta_i)}{f(X/\theta_0)} \times \pi(\theta_i)} = \frac{exp\left[log\left(\frac{f(X/\theta_i)}{f(X/\theta_0)}\right) + log\pi(\theta_i)\right]}{\sum_i exp\left[log\left(\frac{f(X/\theta_i)}{f(X/\theta_0)}\right) + log\pi(\theta_i)\right]},$$

si on suppose que les observations sont indépendantes, on écrit :

$$\pi(\theta_i/X) = \frac{exp\left[\sum_i log\left(\frac{f(x_i/\theta_i)}{f(x_i/\theta_0)}\right) + log\pi(\theta_i)\right]}{\sum_i exp\left[\sum_i log\left(\frac{f(x_i/\theta_i)}{f(x_i/\theta_0)}\right) + log\pi(\theta_i)\right]},$$

ainsi, si $n \to \infty$,

$$\lim_{n\to\infty} \frac{1}{n}\sum_i log\left(\frac{f(x_i/\theta_i)}{f(x_i/\theta_0)}\right) = \int log\left(\frac{f(x_i/\theta_i)}{f(x_i/\theta_0)}\right) f(x_i/\theta_0)\, dx$$

$$= -\int log\left(\frac{f(x_i/\theta_0)}{f(x_i/\theta_i)}\right) f(x_i/\theta_0)\, dx$$

$$= E^{f(x_i/\theta_0)}\left(log\left(\frac{f(x_i/\theta_0)}{f(x_i/\theta_i)}\right)\right) = \begin{cases} 0, & \forall\, \theta_i = \theta_0 \\ -\infty, & \forall\, \theta_i \neq \theta_0 \end{cases},$$

qui est la distance de Kulback-Leibner entre $f(x_i/\theta_i)$ et $f(x_i/\theta_0)$. Donc :

$$\lim_{n\to\infty} \pi(\theta_i/X) = lim \frac{exp\left[\sum_i log\left(\frac{f(x_i/\theta_i)}{f(x_i/\theta_0)}\right) + log\pi(\theta_i)\right]}{\sum_i exp\left[\sum_i log\left(\frac{f(x_i/\theta_i)}{f(x_i/\theta_0)}\right) + log\pi(\theta_i)\right]} = 0, \forall\, \theta_i \neq \theta_0\,, \quad (6)$$

cette dernière équation nous indique que la distribution a postériori devient plus et plus concentrée autour de θ_0, pour $n \to \infty$.

b) Le cas continue

Soit un modèle statistique bayésien caractérisé par un paramètre θ unidmensionnel. En prenant le logarithme des deux membres :

$$ln\, \pi(\theta/x) = ln\, f(x/\theta) + ln\, \pi(\theta) + cte,$$

[2] Un estimateur est convergent de θ_0s'il converge en probabilité.

supposons que la densité a posteriori de θ est unimodale. Un développement de Taylor au voisinage de son mode, disons θ_x^* , donné par :

$$ln\,\pi(\theta/x) = ln\,\pi(\theta_x^*/x) + (\theta - \theta_x^*)\frac{\partial}{\partial\theta}ln\,\pi(\theta/x)|_{\theta=\theta_x^*}$$
$$+\frac{1}{2}(\theta - \theta_x^*)^2\frac{\partial^2}{\partial\theta^2}ln\,\pi(\theta/x)|_{\theta=\theta_x^*} + \cdots,$$

pour $|\theta - \theta_x^*|$ petit, on peut négliger les termes d'ordre supérieur et en tenant compte de la règle de Bayes :

$$ln\,\pi(\theta/x) \cong ln\,f(x/\theta_x^*) + (\theta - \theta_x^*)\frac{\partial}{\partial\theta}ln\,f(x/\theta) + \frac{\partial}{\partial\theta}ln(\pi(\theta))|_{\theta=\theta_x^*}$$
$$+\frac{1}{2}(\theta - \theta_x^*)^2\frac{\partial^2}{\partial\theta^2}ln\,f(x/\theta) + \frac{\partial^2}{\partial\theta^2}ln(\pi(\theta))|_{\theta=\theta_x^*} + \underbrace{ln\,\pi(\theta_x^*) + cte}_{cte},$$

si la densité a priori de θ est « plate » au voisinage du mode θ_x^*, sa dérivée première est nulle en ce point (les dérivées supérieures aussi):

$$ln\,\pi(\theta/x) \cong ln\,f(x/\theta_x^*) + (\theta - \theta_x^*)\frac{\partial}{\partial\theta}ln\,f(x/\theta)|_{\theta=\theta_x^*}$$
$$+\frac{1}{2}(\theta - \theta_x^*)^2\frac{\partial^2}{\partial\theta^2}ln\,f(x/\theta)|_{\theta=\theta_x^*} + cte,$$

dans le cadre asymptotique$(n \to \infty)$, l'influence du prior sur la densité a posteriori est très faible et le mode de la vraisemblance, disons $\hat{\theta}_x$, se confond avec le mode du posterior θ_x^*:

$$ln\,\pi(\theta/x) \cong ln\,f\left(x/\hat{\theta}_x\right) + \frac{1}{2}\left(\theta - \hat{\theta}_x\right)^2\frac{\partial^2}{\partial\theta^2}ln\,f(x/\theta)|_{\theta=\hat{\theta}_x} + cte,$$

le terme impliquant la dérivée première a disparu, car $\hat{\theta}_x$ est le mode de la vraisemblance $\left(\frac{\partial}{\partial\theta}ln\,f(x/\theta)|_{\theta=\hat{\theta}_x} = 0\right)$.

En posant :

$$P(\theta) = -\frac{\partial^2}{\partial\theta^2}ln\,f(x/\theta),$$

on obtient :

$$ln\,\pi(\theta/x) \cong ln\,f\left(x/\hat{\theta}_x\right) - \frac{P(\hat{\theta}_x)}{2}\left(\theta - \hat{\theta}_x\right)^2 + cte,$$

le terme $P\left(\hat{\theta}_x\right)$ ne dépend que des données. On simplifie le script en le notant $\hat{P}_x$.

Pour revenir à la densité a posteriori de θ, il suffit de prendre l'exponentielle des deux membres :

$$\pi(\theta/x) \cong cte \times exp\left(-\frac{\hat{P}_x}{2}\left(\theta - \hat{\theta}_x\right)^2\right)$$

On reconnait la signature fonctionnelle d'une densité normale, localisée sur $\hat{\theta}_x$ et de precision $P(\hat{\theta}_x)$ *:*

$$\theta/x \sim \mathcal{N}\left(\hat{\theta}_x, \left[-\frac{\partial^2}{\partial\theta^2}ln\,f\left(x/\hat{\theta}_x\right)\right]^{-1}\right), \qquad (7)$$

Cette approximation n'est valable que sous les hypothèses : n grand, un seul mode et prior plat au voisinage de celui-ci (voir Boreux et al, 2010).

2. La consistance

La définition d'un estimateur consistant (la convergence au sens de Fisher (Fisher, 1921)) est plus forte que celle d'un estimateur convergent. bragimov et Hasminskii (1981) démontrent que les estimateurs de bayes sont convergents dans un cadre général, c’est-à-dire qu’ils convergent presque sûrement vers la vraie valeur du paramètre lorsque le nombre d’observations tend vers l’infini.

Théorème 1. (bragimov et Hasminskii, (1981))

On pose Θ un ensemble ouvert et borné dans IR^k et la densité $f(x;\theta)$ satisfie les conditionns suivantes :

- $\inf_{|\theta-\theta'|>\gamma} \int_\chi \left(f^{1/2}(x;\theta) - f^{1/2}(x;\theta')\right)^2 dv = k_\theta(\gamma)\ \forall\theta \in \Theta, \gamma > 0$
- $\int_\chi \left(f^{1/2}(x;\theta+h) - f^{1/2}(x;\theta)\right)^2 dv = O\left(\frac{1}{(\ln h)^2}\right), h \to 0, \forall\theta \in \Theta$

Ensuite, l'estimateur $\tilde{t}_n$, qui est bayésien par rapport à la fonction de perte $L(u;\theta) = |u-\theta|^\alpha, \alpha > 0$ et la densité a priori $\pi(\theta)$ lorsque $\pi(\theta)$ est continue et positive en Θ, est un estimateur consistant pour θ.

Dans un cadre non paramétrique, la consistance n'est pas automatique, mais elle dépend de la topologie choisie sur le modèle.

V. CONCLUSION

La statistique bayésienne (ou plus exactement néo-bayésienne vu les deux siècles de décalage entre Bayes et ceux qui s'en réclament actuellement) est une approche de la statistique inférentielle qui propose une réponse unitaire et globale. La mise en œuvre des principes bayésiens, en dehors de cas d'école, s'est longtemps heurtée aux difficultés pratiques de calcul (la distribution a postériori est d’écriture complexe et donc difficilement calculable analytiquement). Aujourd’hui les PC rapides ont modifié la donne, puisque le développement d'algorithmes itératifs possédant des propriétés markoviennes a permis de surmonter cet obstacle de complexité. Cet essai a tendance à se concentrer davantage sur la perspicacité et la compréhension. C'est un guide d'utilisation de l'économétrie bayésienne, pas un traité sur les statistiques mathématiques du sujet. Espérons que les lecteurs finiront l'article avec une meilleure connaissance des méthodes bayésiennes, une compréhension simple et intuitive des principes impliqués et une volonté d'appliquer les techniques bayésiennes à des problèmes empiriques du monde réel.

REFERENCES

[1] Basturk, N, (2013), Historical Developments in Bayesian Econometrics after Cowles Foundation Monographs 10, 14. Tinbergen Institute Discussion Paper 191/III

[2] Bernardo, J. et Smith, A, (1994), Bayesian Theory. John Wiley, New York.

[3] Carlin, B. et Louis, T, (2001), Bayes and Empirical Bayes Methods for Data Analysis. Chapman and Hall, New York, seconde edition.

[4] Gelman, A., Carlin, J., Stern, H., et Rubin, D, (2003), Bayesian Data Analysis. Chapman and Hall, New York, second ´edition.

[5] Geweke, J, (2005), Contemporary Bayesian econometrics and statistics (Vol. 537). John Wiley & Sons.

[6] Ibragimov, I. A., & Hasminskii, R. Z, (1981), Statistical estimation, volume 16 of Applications of Mathematics.

[7] Koop, G, (2003), Bayesian econometrics. John Wiley & Sons.

[8] Koop, G., Poirier, D. J., & Tobias, J. L, (2007), Bayesian econometric methods. Cambridge University Press.

[9] O’Hagan, A, (1994), Bayesian Inference. Numero 2B dans Kendall’s Advanced Theory of Statistics. Chapman and Hall, New York.

[10] O'Hagan, A. et Forster, J, (2002), Bayesian Inference.Numero2Bdans Kendall's Advanced Theory of Statistics. Chapman and Hall, New York, seconde ´edition.

[11] Schervish, M, (1995), Theory of Statistics. Springer-Verlag, New York.