

MINISTERE DE L'ENSEIGNEMENT SUPERIEUR ET DE LA RECHERCHE
SCIENTIFIQUE UNIVERSITE MOULOUD MAMMERI, TIZI OUZOU FACULTE DES
SCIENCES DEPARTEMENT DE MATHEMATIQUES

MEMOIRE DE MAGISTER

SPECIALITE : MATHEMATIQUES
OPTION : PROBABILITES ET STATISTIQUE
Présenté par : Mme HALIL Née KHORSI Rachida

Sujet :

INFERENCE BAYESIENNE EN SERIES CHRONOLOGIQUES

DEVANT LE JURY COMPOSE DE :

Monsieur AIDENE Mohamed, Professeur, U.M.M.T.O	:Président
Melle NOUALI Karima, Maitre de conférences (A), U.M.M.T.O	:Rapporteur
Madame KHELLAS Fazia, Maitre de conférences (A), U.M.M.T.O	: Examinatrice
Monsieur HAMAZ Abdelghani, Maitre de conférences (B), U.M.M.T.O	:Examineur
Madame MERAKEB Farida, Chargée de recherche, U.M.M.T.O	: Invité

Soutenu le 10/03/2011

REMERCIEMENTS

Au terme de ce travail, je tiens tout particulièrement à exprimer ma vive reconnaissance et ma gratitude à Mademoiselle Nouali Karima, Maitre de Conférences à L'U.M.M.T.O, pour avoir proposé et dirigé ce travail. Je la remercie pour sa disponibilité permanente et son soutien.

J'adresse mes vifs remerciements à Monsieur Aidene Mohamed, Professeur à L'U.M.M.T.O, pour l'honneur qu'il me fait en présidant le jury.

Je remercie également Madame Khellas Fazia, Maitre de Conférences à L.M.M.M.T.O, Monsieur Hamaz Abdelghani, Maitre de Conférences à L'U.M.M.T.O ainsi que madame Merakeb Farida, Chargée de Recherche L'U.M.M.T.O pour avoir accepté de lire ce mémoire et de faire partie du jury.

Madame HALIL Ne KHORSI Rachida

Table des Matières

Table des matières	1
Introduction générale	4
Chapitre 1 : Analyse statistique bayésienne.	
1. Introduction	8
2. Principe de l'analyse bayésienne	8
2.1 Distribution a priori	9
2.1.1 Distribution a priori informative	9
2.1.2 Distribution a priori non informative	10
2.2 Distribution a posteriori	13
3. Estimation ponctuelle	14
3.1 Fonction de perte et risque bayésien	15
3.2 Estimateur bayésien	15
4. Tests bayésiens	16
4.1 Cas du test d'hypothèses composites	16
4.2 Cas du test d'une hypothèse ponctuelle	17
5. Réponses au problème de test	17
5.1 Réponses classiques	17
5.1.1 Approche de Neyman-Pearson	17
5.1.2 La p-valeur	18
5.2 Réponses bayésiennes	18
5.2.1 Le facteur de Bayes	21
5.2.2 Odds ratio a posteriori	22
5.2.3 Probabilité a posteriori	22
6. Comparaison des réponses bayésiennes et classiques	24
6.1 Réponse bayésienne la moins favorable à l'hypothèse nulle	25

7. Le paradoxe de Jeffreys-Lindley-Bartlett	27
---	----

Chapitre 2 : Test bayésien de la racine unitaire dans le modèle autorégressif d'ordre 1.

1. Introduction	29
2. Position du problème	30
3. Fonctions de vraisemblance des modèles	31
4. Etude a posteriori	32
4.1 Les densités jointes intégrées des modèles	32
4.2 Les probabilités a posteriori des modèles	33
5. Règle de sélection des modèles	33
6. Les lois a priori du test	33
7. Etude comparative	38
7.1 Les probabilités a posteriori	38
7.2 La puissance des tests	45

Chapitre 3 : Test bayésien d'indépendance dans le modèle autorégressif d'ordre 1.

1. Introduction	49
2. Modélisation de l'autocorrelation	50
3. Position du problème	50
4. Fonctions de vraisemblance des Modèles	51
5. Etude a posteriori	51
5.1 Les densités jointes intégrées des modèles.	51
5.2 Les probabilités a posteriori des modèles	52
6. Les distributions a priori du test	53
7. Etude comparative	55
7.1 Les probabilités a posteriori	55
7.2 La puissance des tests	59

Chapitre 4 : Test bayésien de unit root sous contamination

Première partie : Test bayésien de la racine unitaire en présence de structure de rupture	62
1. Structures de break dans un AR(1)	62
1.1 Changement structurel dans la tendance	63
1.2 Changement structurel dans la moyenne et la variance	63
1.3 Changement structurel dans la moyenne	64
2. Le test de la racine unitaire en présence d'un break	64
2.1 Présentation des modèles	65
2.2 Fonctions de vraisemblance des modèles	66
2.3 Les densités jointes intégrées des modèles	67
3. Etude de Monte-Carlo	68
 Deuxième partie : Test bayésien de la racine unitaire en présence d'outliers additifs	 74
1. Modèle d'outliers en séries chronologiques	75
2. Test de unit root sous contamination AO	75
3. Etude de Monte-Carlo	76
 Conclusion générale	 83
 Bibliographie	 85
 ANNEXE	 91

Introduction générale

Le fondement de la théorie bayésienne s'est basé sur le théorème d'inversion des probabilités connu sous le nom du théorème de Bayes (**Bayes**, 1763). **Laplace** (1774) approfondit la notion de probabilité inverse par l'introduction des probabilités des causes des événements. La démarche de la statistique bayésienne s'est distinguée de celle de la statistique classique (fréquentiste) par le fait qu'elle représente l'incertitude portant sur le paramètre θ du modèle par une distribution de probabilité conditionnellement à un état de connaissance de l'analyste (**Berger**, 1980).

Cette incertitude doit être distinguée de l'aléat qui modélise l'écart du modèle à la réalité. En statistique classique, c'est uniquement ce dernier qui est pris en compte dans la caractérisation du modèle. La représentation de l'incertitude portant sur le paramètre θ du modèle est nommée *loi a priori* de θ . Elle s'interprète comme la représentation formelle sous forme probabiliste de la connaissance sur le paramètre détenue par l'analyste, expertise: approche *subjective*.

Toutefois, si l'on souhaite uniquement traiter l'information apportée par les données (i.e: en faisant abstraction de l'information a priori disponible sur le paramètre; approche *objective*); certains auteurs (**Jeffreys**, 1961, **Bernardo**, 1979, 1980, ...) se sont préoccupés de la construction de lois a priori non informatives.

Par ailleurs, on peut constater que les deux théories, bayésienne et classique possèdent des similitudes dans le cadre de l'estimation. En général, l'estimateur du maximum de vraisemblance coïncide avec l'estimateur bayésien associé à une loi impropre. Par contre les tests, en particulier les tests bilatéraux constituent l'une des situations inférentielles dans lesquelles les réponses bayésiennes et classiques diffèrent considérablement et semblent irréconciliables (**Berger et Sellke**, 1987, **Berger et Delampady**, 1987, **Berger et Yang**, 1994,...).

Dans notre travail, nous portons un intérêt particulier aux tests bayésiens dans les séries chronologiques plus exactement le test de la racine unitaire et le test d'indépendance dans

le modèle AR(1).

Le test d'indépendance a été abordé uniquement par l'approche classique. On rencontre les travaux de **Durbin et Watson** (1950), **Kulberger et Lockhart** (1995), **Andrews et al** (1998) et **Berkoun et al** (2003).

Par ailleurs, la littérature sur les tests bayésiens de la racine unitaire est très volumineuse, tant sur le plan de la théorie statistique que sur celui des applications empiriques et de la théorie économique. **Sims** (1988) fut le premier à aborder le test de la racine unitaire par l'approche bayésienne dans le modèle AR(1). **Dejong et Whiteman** (1991) montrent que la robustesse des tests bayésiens dépend du choix approprié de la loi a priori utilisée. Une attention particulière est donc attribuée aux choix de la loi a priori attribuée au paramètre du modèle.

Ainsi, **Schotman et Van Dijk** (1991) développent un test basé sur les posterior odds. Ils utilisent la loi a priori non informative uniforme pour le paramètre ρ du modèle AR(1). **Berger et Yang** (1994) considèrent la loi a priori non informative de référence. **Lubrano** (1995) utilise la loi Beta symétrique. **Marriott et Newbold** (1998) choisissent la loi a priori Beta non symétrique qui assigne la plus grande masse de probabilité au point $\rho = 1$. **Conigliani et Spezzaferrì** (2007) proposent comme loi a priori, la distribution triangulaire.

On retrouve également dans la littérature actuelle un bon nombre de travaux qui traitent les tests de la racine unitaire avec possibilité de rupture dans la structure du modèle (rupture dans la tendance, dans la moyenne et dans la variance). A titre d'exemple, on peut citer les travaux de **Lumsdaine et Papell** (1997), **Wang et Zivot** (2000) qui considèrent le test de la racine unitaire en présence de break dans la tendance du processus. D'autre part, **Marriott et Newbold**(2000) traitent le cas de présence de break dans la moyenne du processus.

Par ailleurs, certains événements tels que les guerres et les politiques monétaires peuvent affecter les différentes variables macro-économiques. Ainsi la prise en compte des ruptures (breaks) structurelles dans les études empiriques utilisant les séries temporelles est devenue une nécessité. Ceci s'explique par le fait que contrairement à **Nelson et Plosser** (1982), qui utilisent le test classique de Dickey-Fuller, **Perron**(1989) montre que l'évidence de l'existence d'une racine unitaire dans plusieurs variables macro-économiques pourrait être due à la présence d'un important changement structurel dans la tendance des séries.

Cet auteur montre que dans le cas où une série présente une évolution dans sa tendance, les tests habituels de la racine unitaire sont biaisés en faveur de l'hypothèse nulle de racine unitaire. Les résultats de **Perron** (1989) ont été à l'origine de nombreuses recherches en économétrie des séries temporelles qui tiennent compte de cette éventualité de rupture dans la tendance d'une série.

Ce mémoire est organisé comme suit:

Dans le premier chapitre, nous présentons quelques définitions et propriétés des éléments théoriques de l'analyse bayésienne, puis nous abordons les tests d'hypothèses comme un problème de décision. Ce chapitre comprend également une représentation des réponses classiques et bayésiennes aux différents types de tests sous l'angle de la théorie de la décision.

Dans le deuxième chapitre, nous abordons le test bayésien de la racine unitaire dans le modèle autorégressif d'ordre 1. Nous élargissons l'étude de Marriott et Newbold (1998) en introduisant une nouvelle loi a priori et en considérant le cas des petits échantillons. Ainsi une étude comparative au sens des probabilités a posteriori des tests étudiés est effectuée. Nous établissons dans la deuxième partie de ce chapitre une étude comparative des tests bayésiens étudiés avec certains tests classiques au moyen de la puissance.

Dans le troisième chapitre, nous construisons un test bayésien d'indépendance pour le paramètre du modèle AR(1) où différentes lois a priori pour le paramètre du modèle sont proposées.

Dans la deuxième partie de ce chapitre, nous effectuons une étude comparative entre les tests bayésiens proposés et certains tests classiques au moyen de la puissance.

Le dernier chapitre est consacré à l'étude de l'impact de deux types de contamination sur les conclusions du test bayésien de la racine unitaire dans le modèle autorégressif d'ordre 1. Ainsi, dans la première partie de ce chapitre, nous abordons le test en présence d'une rupture au sens de la moyenne. La seconde partie traite le test en présence d'outliers additifs.

Nous terminons par une conclusion générale sur les résultats obtenus en proposant quelques perspectives de recherche.

Abréviations.

UPP(SB) : Uniformement plus puissant (sans biais).

i.i.d : Indépendantes et identiquement distribuées.

$P(M_1/W) > 0.5$: Proportion d'échantillons pour lesquels le modèle M_1 est choisi.

$U[a, b]$: Loi uniforme sur l'intervalle $[a, b]$.

N_{Tr} : Loi normale tronquée.

Beta1 : Loi Beta non symétrique sur $[-1, 1]$ de paramètres $\beta = 0.5$ et $\alpha = 5$.

Beta2 : Loi Beta non symétrique sur $[-1, 1]$ de paramètres $\beta = 0.5$ et $\alpha = 50$.

T1Mexp1 : Loi exponentielle modifiée définie sur $[-1, 1]$ de paramètres $\lambda = 5$ qui attribue une grande masse de probabilité au point 1.

T1Mexp2 : Loi exponentielle modifiée définie sur $[-1, 1]$ de paramètres $\lambda = 50$ qui attribue une grande masse de probabilité au point 1.

T2Mexp1 : Loi exponentielle modifiée définie sur $[-1, 1]$ de paramètres $\lambda = 5$ qui attribue une grande masse de probabilité au point 0.

T2Mexp2 : Loi exponentielle modifiée définie sur $[-1, 1]$ de paramètres $\lambda = 10$ qui attribue une grande masse de probabilité au point 0.

π_β : Puissance de test.

\propto : proportionnel à.

Chapitre 1:

Analyse statistique bayésienne

1 Introduction.

La statistique bayésienne est une théorie concurrente à la statistique dite classique en ce sens que chacune d'elles propose vis-à-vis d'un même problème une approche et une résolution complètement différentes. Nous présentons donc très succinctement les fondements de cette théorie dans une première partie. La deuxième et la troisième partie de ce chapitre sont consacrées à l'inférence bayésienne (test et estimation) sur le paramètre d'un modèle statistique. Dans ce chapitre nous portons un intérêt particulier aux tests bayésiens.

2 Principe de l'analyse bayésienne.

Soit $X = (X_1, \dots, X_n)$ un n -échantillon de variables aléatoires *i.i.d* de densité $f(x/\theta)$ qui dépend d'un paramètre θ inconnu. La vraisemblance des observations x_1, \dots, x_n est notée $L(\underline{x}/\theta)$ où \underline{x} désigne le n -uplet (x_1, \dots, x_n) .

En statistique fréquentiste, on considère θ un paramètre fixe appartenant à un espace $\bar{\Theta}$ et on l'estime sur la base des échantillons observés.

Tandis que dans l'approche bayésienne, on considère θ un paramètre aléatoire et on associe à l'information tirée de l'échantillon, une information prévenant d'une autre source (avis d'analystes, d'experts,...). Cette information additionnelle sur θ est résumée par une loi de probabilité $\pi(\cdot)$ dite "*loi a priori*" du paramètre θ .

L'inférence bayésienne est alors fondée sur une loi dite "*loi a posteriori*" de θ qui est une fonction de vraisemblance du modèle et de la loi a priori de θ .

Définition 2.1 *Un modèle statistique bayésien est la double donnée d'un modèle paramétrique $\{f_\theta(x), \theta \in \overline{\Theta}\}$ et une loi de probabilité, de densité π , dite loi a priori qui est la loi marginale de la variable aléatoire Θ .*

2.1 Distribution a priori.

Le choix de la loi a priori du paramètre du modèle est crucial pour l'analyse bayésienne. Cette première touche d'une façon directe les distributions a posteriori qui sont utilisées dans l'inférence bayésienne.

Il y'a deux modes de pensées qui sont misent à l'evidence par rapport a la loi a priori.

Le premier est "*subjectif*". Il repose sur l'information disponible sur le paramètre obtenue des opinions des chercheurs et d'experts. Cette information est exprimée par une loi de probabilité dite "distribution a priori informative".

Cependant, la théorie bayésienne peut être appliquée même dans le cas où on ne dispose pas d'informations a priori; c'est le deuxième mode de pensées qui est plutôt "*objectif*". Là encore, le paramètre est considéré comme une valeur d'une variable aléatoire et on lui assigne une loi de probabilité dite "distribution a priori non informative".

2.1.1 Distribution a priori informative.

La modélisation a priori informative est le point le plus délicat de l'analyse bayésienne. Il existe plusieurs procédés pour obtenir des lois informatives (voir Dreesbeke et al 2002). Nous donnons la description de l'un des plus intéressants des procédés qui est celui des familles naturelles conjuguées.

Familles naturelles conjuguées.

Définition 2.2 *Une famille G de lois a priori pour θ est dite conjuguée pour la vraisemblance $L(\underline{x}/\theta)$ si pour tout $\pi_\theta(.) \in G$, la loi a posteriori $\pi(\theta/X)$ est un élément de G .*

L'approche conjuguée reste la solution la plus standard dans le cadre informatif. Une famille naturelle conjuguée est un concept qui simplifie considérablement le calcul des distributions a posteriori. C'est donc la raison du développement de ce type de lois a priori.

Il faut noter que seuls les modèles à structure exponentielle admettent une famille conjuguée.

Exemple 1.

On donne dans la table 1 quelques lois a priori conjuguées.

Table 1 : Lois a priori conjuguées.

$f(x/\theta)$	$\pi(\theta)$	$\pi(\theta/X)$
Binomiale $B(n, \theta)$	Beta $Beta(\alpha, \beta)$	Beta $Beta(\alpha + x, \alpha + n - x)$
Normale $N(\theta, \sigma^2)$	$N(\mu, \tau^2)$	$N(\varphi(\sigma^2\mu + \tau^2x), \varphi\sigma^2\tau^2)$ avec $\varphi^{-1} = \sigma^2 + \tau^2$
Poisson $P(\theta)$	Gamma $G(\alpha, \beta)$	Gamma $G(\alpha + x, \beta + 1)$
Gamma $G(\nu, \theta)$	gamma $G(\alpha, \beta)$	gamma $G(\alpha + \nu, \beta + x)$

2.1.2 Distribution a priori non informative.

Ces distributions sont conçues dans le but de faire de l'analyse bayésienne lorsqu'il y a absence de l'information a priori sur le paramètre d'intérêt ou dans le cas où c'est difficile de traduire en terme de loi a priori l'information disponible sur les paramètres par une loi de probabilité .

Le choix d'une loi a priori non informative conduit souvent à la spécification d'une mesure et non d'une probabilité, c-à-d une loi impropre (approche dite objective).

Une loi impropre est une loi qui est définie sur l'espace \mathbf{R} de densité π non intégrable.

$$\int_{\mathbf{R}} \pi(\theta) d\theta = +\infty$$

Les lois a priori impropres forment une extension nécessaire des lois a priori propres. Dans un contexte non informatif, elles permettent en général une meilleure modélisation de l'absence d'information a priori, notamment lorsque l'espace des paramètres n'est pas compact. A l'exemple de la loi uniforme a priori $\pi(\theta) = 1$ sur \mathbf{R} qui n'est autre que la mesure de lebesgue sur \mathbf{R} considérée comme une extension de loi a priori uniforme définie sur un compact.

Nous exposons brièvement les trois approches:

La recherche d'a priori invariante, l'approche de Jeffreys et l'approche dite "a priori de référence".

Une analyse objective plus convaincante repose sur la construction des lois a priori qui respecte une invariance complète (loi de Jeffreys) ou au moins partielle (pour un sous groupe donné de transformations du paramètre, par exemple les translations).

a) Mesures a priori invariantes:

Définition 2.3 (Invariance par transformation sur le paramètre).

Une mesure M sur $\bar{\Theta}$ est invariante par transformation h de $\bar{\Theta}$ dans $\bar{\Theta}$ si M est identique à son image par h définie par Moh^{-1} .

Remarque 1.

Si M est caractérisée par sa densité π et h est bijective bidérivable, la densité de Moh^{-1} est égale à $\|\rho^{-1}\| (Moh^{-1})$, la condition d'invariance s'exprime par l'égalité:

$$\pi = \|\partial h^{-1}\| (\pi o h^{-1})$$

où $\|\partial \rho^{-1}\|$ est la valeur absolue du déterminant de la matrice des dérivées partielles.

Exemple 2. (Berger et Yang (1995)).

Dans le modèle AR(1), en prenant la transformation:

$$h : \rho \longmapsto h(\rho) = \frac{1}{\rho}, \quad |\rho| > 1$$

et la loi a priori:

$$\pi_{SR}(\rho) = \begin{cases} 1/(2\pi\sqrt{1-\rho^2}) & \text{si } |\rho| < 1 \\ 1/(2\pi |\rho| \sqrt{\rho^2-1}) & \text{si } |\rho| > 1 \end{cases}$$

Dans ce cas, la condition d'invariance par la transformation h sur le paramètre ρ s'exprime par l'égalité

$$\pi(h(\rho)) = \pi(\rho) \left| \frac{\partial h(\rho)}{\partial \rho} \right|^{-1}$$

Définition 2.4 (Invariance par rapport à une statistique suffisante)

Si T est une statistique suffisante pour le modèle $L(\underline{x}/\theta)$ alors l'a posteriori non subjective $\pi(\theta/X)$ obtenue pour ce modèle est la même que l'a posteriori non subjective $\pi(\theta/t)$ obtenue pour le modèle $L(t/\theta)$.

Pour plus de détails sur le rôle de la propriété d'invariance dans la sélection de la loi a priori non subjective, on peut consulter Hartigan (1964), Davis (1983), Dreesbeke et al (2002).

b) Mesure a priori de Jeffreys.

Le mode de spécification d'une loi a priori non informative connue sous le nom de la mesure a priori de Jeffreys consiste à assigner à un modèle d'échantillonnage caractérisé par sa vraisemblance $L(\underline{x}/\theta)$, la mesure a priori de densité:

$$\pi_J(\theta) = |\det I(\theta)|^{1/2}$$

où θ est un vecteur de paramètres inconnus et $I(\theta)$ est la matrice d'information de Fischer définie par:

$$I(\theta) = I_{ij}(\theta) = -E\left[\frac{\partial^2}{\partial\theta_i\partial\theta_j} \log L(\underline{x}/\theta)\right]$$

Cette mesure a priori est invariante par transformation bijective.

Exemple 3. (Berger et Yang (1995)).

Soit $X \sim N(\mu, \sigma^2)$ où la moyenne μ et la variance σ^2 sont des paramètres inconnus.

La fonction de densité s'écrit:

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left\{-\frac{1}{2\sigma^2}(x - \mu)^2\right\}$$

Le paramètre d'intérêt est le vecteur $\theta = (\mu, \sigma)$.

La matrice d'information de Fisher est donnée comme suit

$$I(\theta) = \begin{pmatrix} \frac{1}{\sigma^2} & 0 \\ 0 & \frac{2}{\sigma^2} \end{pmatrix}$$

La loi a priori de Jeffreys s'écrit

$$\pi_J(\theta) \propto \frac{1}{\sigma^2}$$

c) Lois a priori de référence.

La théorie de l'a priori de référence est introduite par Bernardo (1979) suite à la difficulté rencontrée dans l'utilisation de l'approche de Jeffreys dans le cas multidimensionnel. C'est un procédé qui consiste à assigner comme loi de probabilité a priori, une densité $\pi^*(\theta)$ qui maximise l'information apportée par les données sur le paramètre θ lorsque la loi a priori

est $\pi(\theta)$ qui est définie par (Lindley, 1956) :

$$I_X^\theta(\pi) = E^X(D(\pi(\theta/X), \pi(\theta))) = \int D(\pi(\theta/X), \pi(\theta))m(x)dx \quad (1)$$

où $m(x) = \int_{\Theta} L(\underline{x}/\theta)\pi(\theta)d\theta$, représente la densité marginale de l'échantillon X .

$D(\pi(\theta/X), \pi(\theta))$ désigne la divergence de Kullback-Leibler entre la loi a priori et la loi a posteriori qui est définie par:

$$D(\pi(\theta/X), \pi(\theta)) = \int \pi(\theta/X) \log \frac{\pi(\theta/X)}{\pi(\theta)} d\theta$$

l'a priori $\pi^*(\theta)$ qui maximise la quantité $I_X^\theta(\pi)$ vérifie l'équation:

$$\pi(\theta) \propto \exp\left\{\int L(\underline{x}/\theta) \log \pi(\theta/X) d\theta\right\}. \quad (2)$$

En général, la solution exacte de l'équation (2) est difficile à obtenir. Cependant, une solution approximée peut être obtenue pour de grands échantillons.

2.2 Distribution a posteriori.

Soit $X = (X_1, \dots, X_n)$ un n-échantillon de variables aléatoires i.i.d de densité $f(x/\theta)$. On suppose que θ est une valeur d'une variable aléatoire Θ de densité $\pi(\theta)$ et soit

$$L(\underline{x}/\theta) = \prod_{i=1}^n f(x_i/\theta),$$

la fonction de vraisemblance des observations.

La distribution a posteriori de θ notée $\pi(\theta/\underline{x})$ représente la distribution conditionnelle de θ sachant les observations $\underline{x} = (x_1, \dots, x_n)$.

$$\pi(\theta/\underline{x}) = \frac{L(\underline{x}/\theta)\pi(\theta)}{\int_{\underline{\Theta}} L(\underline{x}/\theta)\pi(\theta)d\theta} = \frac{\prod_{i=1}^n f(x_i/\theta)\pi(\theta)}{\int_{\underline{\Theta}} \prod_{i=1}^n f(x_i/\theta)\pi(\theta)d\theta}$$

où encore, on peut écrire:

$$\pi(\theta/\underline{x}) \propto \prod_{i=1}^n f(x_i/\theta)\pi(\theta)$$

où \propto signifie proportionnel à.

La distribution a posteriori $\pi(\theta/\underline{x})$ donne alors l'information dont on dispose sur θ , après observation. Elle représente un compromis entre l'information a priori (donnée par π) et l'information tirée de l'échantillon \underline{x} (donné par $L(\underline{x}/\theta)$).

3 Estimation ponctuelle.

3.1 Fonction de perte et risque bayésien.

a) Fonction de perte.

Définition 3.1 Soit $T = T(x_1, \dots, x_n)$, un estimateur de θ . On appelle "fonction de perte" et on note $l(t, \theta)$, toute fonction satisfaisant:

- $l(t, \theta) \geq 0$ pour toute valeur d'estimateur t et $\forall \theta \in \bar{\Theta}$.
- $l(t, \theta) = 0$ si $t = \theta$.

Cette fonction mesure la perte occasionnée lorsqu'on estime θ par t .

Nous représentons quelques fonctions de perte rencontrées dans la littérature:

1. $l_1(t, \theta) = c_1(t - \theta)1_{\{\theta \leq t\}}(t) + c_2(\theta - t)1_{\{\theta > t\}}(t)$.
2. $l_2(t, \theta) = \varphi(\theta)|t - \theta|^r$ avec $\varphi(\theta) \geq 0$ et $r > 0$.
3. $l_3(t, \theta) = \begin{cases} A & \text{si } |t - \theta| > \epsilon \quad \forall \epsilon > 0, \quad A > 0 \\ 0 & \text{si } |t - \theta| \leq \epsilon \end{cases}$

Remarques 2.

- i) Si $\varphi(\theta) = 1$, et $r=2$, $l_2(t, \theta) = (t - \theta)^2$ est l'erreur quadratique.
- ii) Si $c_1 = c_2$, $l_1(t, \theta) = |t - \theta|$ est l'erreur absolue.

b) Risque bayésien.

Définition 3.2 Soit $T = T(x_1, \dots, x_n)$ un estimateur de θ et $l(t, \theta)$, la fonction de perte associée. On appelle risque bayésien moyen noté $R_T(\theta)$, la quantité:

$$R_T(\theta) = E_\theta[l(t, \theta)] = \int_{\mathcal{X}} l(t, \theta) L(\underline{x}/\theta) d\underline{x}$$

Définition 3.3 Moyennant la fonction de perte $l(t, \theta)$ associée à un estimateur $T = T(X_1, \dots, X_n)$, on définit par le risque bayésien, le nombre réel

$$\tau(T) = \int_{\bar{\Theta}} R_T(\theta) \pi(\theta) d\theta$$

3.2 Estimateur bayésien.

Soit $X = (X_1, \dots, X_n)$ un n-échantillon de variables aléatoires *i.i.d* de densité $f(x/\theta)$, où θ est un paramètre inconnu à estimer.

En statistique bayésienne, on suppose θ est une valeur d'une variable aléatoire auquel on assigne une loi a priori $\pi(\theta)$, puis on détermine la distribution a posteriori de θ via le théorème de Bayes. L'estimateur de Bayes est alors défini comme suit:

Définition 3.4 *On appelle estimateur de Bayes associé à une fonction de perte $l(t, \theta)$ et à une loi a priori $\pi(\cdot)$, tout estimateur (lorsqu'il existe) qui minimise le risque de Bayes.*

Nous donnons dans la proposition suivante quelques formes d'estimateurs de Bayes.

Proposition 3.1 .

Si θ est un paramètre unidimensionnel et

i) si $l(t, \theta) = \begin{cases} c & \text{si } t \neq \theta \\ 0 & \text{si } t = \theta \end{cases}$ avec $c > 0$, l'estimateur de Bayes correspond au mode de la distribution a posteriori $\pi(\theta/\underline{x})$.

ii) si $l(t, \theta) = c_1(t - \theta)1_{\{\theta \leq t\}}(t) + c_2(\theta - t)1_{\{\theta > t\}}(t)$, l'estimateur de Bayes T^ est le quantile $\frac{c_1}{c_1 + c_2}$ de la distribution a posteriori.*

$$(c.\grave{a}.d. \quad P(\Theta < t^*/X) = \frac{c_1}{c_1 + c_2}).$$

Lorsque $c_1 = c_2$, l'estimateur de Bayes est la médiane a posteriori.

iii) si $l(t, \theta) = (t - \theta)^2$, l'estimateur de Bayes est la moyenne a posteriori: $T^ = E(\theta/X)$.*

Définition 3.5 *Un estimateur T_0 est inadmissible s'il existe un estimateur T_1 qui domine T_0 , c'est-à-dire tel que pour tout θ :*

$$R(\theta, T_0) \geq R(\theta, T_1)$$

et, pour au moins une valeur θ_0 du paramètre,

$$R(\theta_0, T_0) > R(\theta_0, T_1)$$

Définition 3.6 *Un estimateur T est minimax si:*

$$\forall T', \quad \sup_{\theta} R(T, \theta) \leq R(T', \theta)$$

c'est-à-dire si $\sup_{\theta} R(T, \theta) = \inf_{T'} \sup_{\theta} R(T', \theta)$.

4 Tests bayésiens.

Tout test d'hypothèses peut être vu comme un problème d'estimation et traité donc sur un angle décisionnel (voir Robert, 1992, Droesbeke et al, 2002).

4.1 Cas de test d'hypothèses composites.

Soit un échantillon (X_1, \dots, X_n) de densité $f(x/\theta)$, $\theta \in \overline{\Theta}$, dont on désire tester

$$H_0 : \theta \in \overline{\Theta}_0 \quad \text{contre} \quad H_1 : \theta \in \overline{\Theta}_1.$$

Ce test peut être formulé par la recherche d'un estimateur de la fonction indicatrice $1_{\theta_0}(\theta)$ à valeur dans $\{0, 1\}$.

Ainsi, la fonction de coût naturelle associée à un problème de test est le coût '0-1' définie par:

$$L_1(\theta, \delta(\underline{x})) = \begin{cases} 1 & \text{si } \delta(\underline{x}) \neq 1_{\overline{\Theta}_0}(\theta) \\ 0 & \text{sinon} \end{cases} \quad (3)$$

où $\delta(\underline{x})$ est la décision du statisticien lorsque l'observation est \underline{x} .

Quand $\delta(\underline{x}) = 1$, on accepte H_0 et quand $\delta(\underline{x}) = 0$, on refuse H_0 .

L'expression (3) signifie que lorsqu'on choisit la bonne hypothèse, la perte liée à cette décision est nulle. Dans le cas contraire, elle vaut 1. Plus généralement, on peut considérer la fonction de coût dissymétrique suivante

$$L_{a_0-a_1}(\theta, \delta(\underline{x})) = \begin{cases} a_0 & \text{si } \theta \in \overline{\Theta}_0 \quad \text{et} \quad \delta(\underline{x}) = 0, \quad (\text{rejeter } H_0) \\ a_1 & \text{si } \theta \notin \overline{\Theta}_0 \quad \text{et} \quad \delta(\underline{x}) = 1, \quad (\text{accepter } H_0) \\ 0 & \text{si } \delta(\underline{x}) = 1_{\overline{\Theta}_0}(\theta), \quad \text{quad}(\text{la bonne décision}) \end{cases}$$

où a_0 et a_1 sont des réels strictement positifs.

Dans le cas des tests bayésiens, on considère une loi a priori π sous la forme d'un mélange

de lois de type

$$\pi(\theta) = P(H_0) \cdot \pi_0(\theta) + (1 - P(H_0)) \cdot \pi_1(\theta)$$

Ainsi, dans l'approche bayésienne, on attribue à l'hypothèse H_0 une probabilité a priori $P(H_0)$ et on répartit le reste de la masse de probabilité à l'hypothèse H_1 .
(i.e: $P(H_1) = 1 - P(H_0)$) selon une loi de densité π_1 .

4.2 Cas du test d'une hypothèse ponctuelle.

Les tests d'une hypothèse ponctuelle méritent cependant une attention particulière. Si la loi a priori est continue l'événement $\{\theta = \theta_0\}$ est de masse nulle et donc de probabilité (a priori et a posteriori) nulle: $\pi(\theta = \theta_0) = \pi(\theta = \theta_0/\underline{x}) = 0$.

Ce qui veut dire qu'une hypothèse ponctuelle devrait systématiquement être rejetée dans le cas continu.

Cette difficulté a pu être contournée en modifiant la loi a priori π , afin qu'elle affecte une probabilité non nulle à l'événement $\{\theta = \theta_0\}$. La densité de la nouvelle loi a priori $\tilde{\pi}$ s'écrit alors sous la forme:

$$\tilde{\pi}(\theta) = P(H_0).\pi_0(\theta) + (1 - P(H_0)).\pi_1(\theta)$$

avec $\pi_0(\theta) = \delta_{\theta_0}(\theta)$ est la masse de Dirac en θ_0 .

Dans le paragraphe suivant, on représente la règle de décision en fonction de coûts respectifs aux hypothèses nulle et alternative.

5 Réponses au problème de test.

5.1 Réponses classiques.

5.1.1 Approche de Neyman-Pearson.

Dans l'approche de Neyman et Pearson, on fixe le seuil $\alpha = P(\text{rejeter } H_0 / H_0 \text{ vraie})$ et on recherche directement des tests uniformément plus puissants sans biais notés UPPSB (ou uniformément plus puissants notés UPP pour les tests unilatéraux).

Mais plutôt que de fixer α , Fisher (1956) suggère de déterminer la p-valeur (ou niveau exact de significativité) associé au problème de test. Cela implique en particulier pour Fisher, le niveau de test est déterminé après observation de x et non pas avant comme le suggèrent Neyman et Pearson.

5.1.2 La p-valeur.

Définition 5.1 Soit $T(X)$, une statistique du test et $X=x$ est l'échantillon observé. Soit $t=T(x)$, la valeur de la statistique qui correspond à x , alors la p-valeur (ou le seuil de significativité observé) est

$$p(t) = P_{\theta_0}(|T(X)| \geq |t|)$$

La p-valeur est à valeurs dans l'intervalle $[0, 1]$. Sa comparaison avec le seuil α choisi avant le début du test permet de répondre au problème de test. On accepte l'hypothèse nulle H_0 si $p(x)$ est supérieure à α et on la refuse dans le cas contraire.

Critiques faite sur la p-valeur.

La p-valeur prend en compte des événements qui ne se sont pas produit (Jeffreys, 1961) du type $\{T(X) \geq t(x)\}$ alors que le véritable événement est $\{T(X) = t(x)\}$. L'approche classique omet donc de l'information au détriment de l'hypothèse nulle.

Remarques 3.

1. Pour un test bilatéral:

$$p(t) = \begin{cases} 2F_0(t) & \text{si } F_0(t) < 0.5 \\ 2(1 - F_0(t)) & \text{si } F_0(t) \geq 0.5 \end{cases}$$

2. Pour un test unilatéral à droite (rejet des valeurs trop grandes):

$$p(t) = 1 - F_0(t)$$

3. Pour un test unilatéral à gauche (rejet des valeurs trop petites):

$$p(t) = F_0(t)$$

où F_0 représente la fonction de répartition de T sous H_0 .

5.2 Réponses bayésiennes.

5.2.1 Le facteur de Bayes.

Le facteur de Bayes est un critère bayésien de sélection de modèles, comme il est un outil pour comparer la crédibilité de deux hypothèses.

a) Cas de sélection des modèles.

Lorsqu'on est appelé à faire une sélection entre deux modèles M_1 et M_2 par l'approche bayésienne, on affecte à chacun des deux modèles des probabilité a priori $P(M_i)$, $i = 1, 2$ vérifiant

$$P(M_1) + P(M_2) = 1.$$

Du théorème de Bayes, on peut alors exprimer les probabilités a posteriori

$$P(M_i/X) = \frac{P(X/M_i)P(M_i)}{m(X)}, \quad i = 1, 2.$$

où $X = (X_1, \dots, X_n)$ est le n-échantillon, et $m(X) = P(X/M_1)P(M_1) + P(X/M_2)P(M_2)$, représente la probabilité marginale des données (qu'on suppose non nulle).
avec $P(M_2/X) = 1 - P(M_1/X)$.

On peut écrire

$$\frac{P(M_1/X)}{P(M_2/X)} = \frac{P(X/M_1)}{P(X/M_2)} \times \frac{P(M_1)}{P(M_2)} \quad (4)$$

où

$$B_{12} = \frac{P(X/M_1)}{P(X/M_2)}$$

représente le facteur de Bayes en faveur du modèle M_1 contre le modèle M_2 .

Le rapport $\frac{P(M_1/X)}{P(M_2/X)}$ est dit quotient d'enjeux a posteriori en faveur du modèle M_1 contre le modèle M_2 . Dans la littérature anglaise il est dit "posterior odds ratio".

Le rapport $\frac{P(M_1)}{P(M_2)}$ est dit quotient d'enjeux a priori en faveur de M_1 contre M_2 . Dans la littérature anglaise, il est dit "prior odds ratio".

On peut alors exprimer la relation (4) comme suit:

$$\text{Quotient d'enjeux a posteriori} = \text{Facteur de Bayes} \times \text{Quotient d'enjeux a priori}.$$

Remarque 4. Si on accorde le même poids a priori pour les deux hypothèses i.e:

$$P(M_1) = P(M_2) = \frac{1}{2}.$$

Le facteur de Bayes s'écrit

$$B_{12} = \frac{P(M_1/X)}{P(M_2/X)}.$$

Dans cette situation, il correspond exactement au quotient d'enjeux a posteriori.

b) Cas de test d'hypothèses.

On considère le test d'hypothèses composites:

$$H_0 : \theta \in \underline{\Theta}_0 \quad \text{contre} \quad H_1 : \theta \in \overline{\Theta}_1$$

La loi a priori est donnée par le mélange de lois

$$\pi(\theta) = P(H_0)\pi_0(\theta) + (1 - P(H_0))\pi_1(\theta)$$

où $\pi_i(\theta) = P(\theta/H_i)$ est la probabilité a priori de θ sous l'hypothèse H_i .
et $P(H_1) = 1 - P(H_0)$.

La fonction de vraisemblance sous H_i s'écrit:

$$L(X/H_i) = \int_{\theta \in \overline{\Theta}_i} L(X/H_i, \theta) \pi_i(\theta) d\theta, \quad i = 0, 1$$

Le facteur de Bayes de l'hypothèse H_0 en faveur de l'hypothèse H_1 s'exprime par:

$$B_{01} = \frac{L(X/H_0)}{L(X/H_1)}.$$

Comme nous l'avons déjà indiqué, le facteur de Bayes donne un indicateur objectif de l'hypothèse $H_0 : \theta \in \underline{\Theta}_0$. Malheureusement, l'utilisation d'une loi impropre rend impossible le calcul de ce facteur. En effet, les probabilités a posteriori des hypothèses nulle et alternatives ne sont pas définies pour une telle loi. La résolution générale de cette incompatibilité entre les tests bayésiens et lois a priori impropres reste un problème ouvert, même si plusieurs solutions partielles ont été déjà proposées, via la définition de "pseudo-facteur de Bayes" notamment, que nous n'aborderons pas ici (le facteur de Bayes fractionné, voir Conigliani et O'Hagan, 2000 et le facteur de Bayes intrinsèque, voir O'Hagan, 1995, 1997).

Remarques 5.

1. Lorsque les deux hypothèses sont équiprobables, le facteur de Bayes correspond au rapport des probabilités a posteriori de H_0 sur H_1 .

$$B_{01} = \frac{P(H_0/X)}{P(H_1/X)}$$

2. Dans le cas d'un test d'une hypothèse simple contre une hypothèse composite suivant:

$$H_0 : \theta = \theta_0 \quad \text{contre} \quad H_1 : \theta \neq \theta_0$$

où $P(H_0) = P(H_1) = \frac{1}{2}$ et $\pi_0(\theta)$ est la masse de Dirac au point θ_0 .

Le facteur de Bayes s'écrit

$$B_{01} = \frac{L(X/\theta_0)}{\int_{\theta \neq \theta_0} L(X/\theta) \pi_1(\theta) d\theta}$$

3. Le facteur de Bayes est une quantité non bornée, il peut prendre des valeurs entre 0 et $+\infty$.

Interprétation:

Kass et Raftery (1995) proposent un guide d'interprétation des valeurs du facteur de Bayes qui peut servir d'aide à la décision pour écarter ou non une hypothèse jugée moins crédible.

Table 2: Guide d'interprétation du facteur de Bayes.

Kass et Raftery (1995)	
B_{10}	H_0
De 1 à 3.2	rejetée.
De 3.2 à 10	fortement rejetée.
De 10 à 100	très fortement rejetée.
>100	rejetée de façon décisive.

5.2.2 Quotient d'enjeux a posteriori.

le quotient d'enjeux a posteriori est un rapport largement utilisé dans les études médicales. C'est un cas particulier du facteur de Bayes. Il correspond au rapport des probabilités a posteriori des deux modèles (posterior odds ratio) qui s'écrit:

$$K = \frac{P(M_1/W)}{P(M_2/W)} = \frac{P(M_1/W)}{1 - P(M_1/W)}$$

Remarque 6.

Le quotient d'enjeux a posteriori est une quantité positive ou nulle.

Interprétation

- Si $K > 1$ alors on choisit le modèle M_1 .
- Si $K < 1$ alors le choix sera porté sur le modèle M_2 .
- $K = 1$ situation d'indécision.

5.2.3 Probabilité a posteriori

Dans le cas d'un problème de sélection entre deux modèles M_1 et M_2 , la probabilité a posteriori du modèle M_i , $i = 1, 2$ s'écrit

$$P(M_i/X) = \frac{P(M_i)P(X/M_i)}{\sum_{j=1}^2 P(M_j)P(X/M_j)}, \quad i = 1, 2$$

qu'on peut formuler en fonction du facteur de Bayes B_{12} ,

$$P(M_1/X) = \frac{1}{1 + \frac{P(M_2)}{P(M_1)} \times \frac{P(X/M_2)}{P(X/M_1)}} = \frac{1}{1 + \frac{P(M_2)}{P(M_1)} B_{12}^{-1}}$$

De même, on déduit

$$P(M_2/X) = \frac{1}{1 + \frac{P(M_1)}{P(M_2)} \times \frac{P(X/M_1)}{P(X/M_2)}} = \frac{1}{1 + \frac{P(M_1)}{P(M_2)} B_{12}}$$

avec $P(M_1) + P(M_2) = 1$

Remarque 7.

Lorsqu'on est devant un test d'une hypothèse H_0 contre une hypothèse H_1 , on peut formuler les différentes probabilités a posteriori $P(H_0/X)$ et $P(H_1/X)$ en suivant le même procédé cité au dessus.

Exemple 3 (Berger et Sellke, 1987).

Soit $X = (X_1, X_2, \dots, X_n)$ un échantillon de variables aléatoires *i.i.d* de loi $N(\theta, \sigma^2)$ avec σ connu. On s'intéresse au test d'hypothèses

$$H_0 : \theta = \theta_0 \quad \text{contre} \quad H_1 : \theta \neq \theta_0$$

La statistique de test s'écrit:

$$t = T(X) = \sqrt{n}(\bar{X} - \theta_0)/\sigma \sim N(0, 1)$$

où $\bar{X} \sim N(\theta_0, \sigma^2/n)$

La p-valeur s'écrit

$$p(t) = 2(1 - \phi(t))$$

avec ϕ est la fonction de répartition de la loi normale $N(0, 1)$.

Si on considère la loi a priori $\pi(\theta) \sim N(\theta, \sigma^2)$ alors la loi marginale de \bar{X} est $N(\theta_0, \sigma^2(1+n^{-1}))$

Dans ce cas le facteur de Bayes s'écrit

$$\begin{aligned} B_{01} &= \frac{(2\pi\sigma^2/n)^{-1/2} \exp\{-\frac{n}{2\sigma^2}(\bar{X} - \theta_0)^2\}}{(2\pi\sigma^2(1+n^{-1}))^{-1/2} \exp\{-\frac{1}{2(\sigma^2(1+n^{-1}))}(\bar{X} - \theta_0)^2\}} \\ &= (1+n)^{1/2} \exp\{-\frac{1}{2(1+n^{-1})}t^2\} \end{aligned}$$

La probabilité a posteriori sous H_0 est

$$P(H_0/X) = [1 + \frac{1 - P(H_0)}{P(H_0)} (1+n)^{-1/2} \exp\{\frac{1}{2(1+n^{-1})}t^2\}]^{-1}$$

Le tableau suivant donne les valeurs de $p(t)$ et la probabilité a posteriori $P(H_0/X)$ avec $P(H_0) = P(H_1) = \frac{1}{2}$.

Table 3. Valeurs de $p(t)$ et $P(H_0/X)$.

t	p(t)	$P(H_0/X)$						
		n=1	n=5	n=10	n=20	n=50	n=100	n=1000
1.645	0.10	0.42	0.44	0.47	0.56	0.65	0.72	0.89
1.960	0.05	0.35	0.33	0.37	0.42	0.52	0.60	0.82
2.576	0.01	0.21	0.13	0.14	0.16	0.22	0.27	0.53
3.291	0.001	0.086	0.028	0.024	0.026	0.034	0.045	0.124

Cet exemple montre que la réponse classique et la réponse bayésienne ne sont pas en accord. En effet si par exemple $n=50$ et $t=1.960$ les bayésiens concluent en faveur de l'hypothèse nulle avec $P(H_0/X) = 0.52$ alors que les fréquentistes rejettent H_0 avec $P(t) = 0.05$.

Il est à noter qu'il existe d'autres critères bayésiens de comparaison de modèles tels que le critère d'information bayésien noté BIC introduit par Schwarz (voir Schwarz, 1978), et le critère d'information de déviance noté DIC (voir Spiegelhalter et al 2002).

6 Comparaison des réponses bayésiennes et classique.

Nous évoquons les trois théorèmes suivants (Robert, 2000, Dreesbeke et al, 2002):

Théoreme 6.1 *Dans l'approche classique, un estimateur de $1_{\theta_0}(\theta)$ admissible et maximax pour le coût $L_{a_0-a_1}(\theta, \delta(\underline{x}))$ est:*

$$\delta(\underline{x}) = \begin{cases} 1 & \text{si } p(\underline{x}) > \frac{a_1}{a_0 + a_1} \\ 0 & \text{sinon} \end{cases}$$

où $p(\underline{x})$ est la p-valeur associée au problème de test.

Dans l'approche classique, on rejette H_0 si la p-valeur est inférieure à $\frac{a_1}{a_0 + a_1}$. Ce qui veut dire qu'on raisonne directement par rapport à la p-valeur.

Théoreme 6.2 *L'analyse bayésienne du test pour le coût $L_{a_0-a_1}$ conduit à l'estimateur*

$$\delta^\pi(\underline{x}) = \begin{cases} 1 & \text{si } P(H_0/X) > \frac{a_1}{a_0 + a_1} \\ 0 & \text{sinon} \end{cases}$$

où $P(H_0/X)$ est la probabilité a posteriori de l'hypothèse nulle et π est la loi a priori considérée.

Un statisticien bayésien rejette l'hypothèse nulle si la probabilité a posteriori est inférieure à $\frac{a_1}{a_0 + a_1}$.

Comme le facteur de Bayes est relié à la probabilité a posteriori, on peut alors donner le résultat suivant.

Théoreme 6.3 *Pour le coût $L_{a_0-a_1}$, l'estimateur de Bayes de $1_{\theta_0}(\theta)$ est*

$$\delta^\pi(\underline{x}) = \begin{cases} 1 & \text{si } B_{01} > \frac{a_1}{a_0} \times \frac{1 - P(H_0)}{P(H_0)} \\ 0 & \text{sinon} \end{cases}$$

Ici, le seuil d'acceptation de l'hypothèse H_0 dépend du rapport $\frac{a_1}{a_0}$ et du rapport a priori des deux hypothèses $\frac{1 - P(H_0)}{P(H_0)}$.

Comme l'indique le théorème 6.1 et théorème 6.2, la réponse au problème de test dépend de la p-valeur dans l'approche classique ou de la probabilité a posteriori pour les bayésiens par rapport à un même seuil fonction du coût choisi.

Ainsi, la comparaison des réponses classiques et bayésiennes en théorie des tests revient donc à la comparaison directe de la p-valeur et de la probabilité a posteriori.

La probabilité a posteriori de l'hypothèse nulle et le facteur de Bayes constituent des réponses au problème de test bayésien. Ces dernières sont donc souvent comparées à la p-valeur qui représente la réponse classique.

Comme les réponses bayésiennes dépendent fortement de la loi a priori choisie, on peut considérer au lieu d'une seule loi a priori, une classe de lois a priori.

6.1 Réponse bayésienne la moins favorable à l'hypothèse nulle.

Berger et Sellke (1987) et Berger et Delampady (1987) ont adopté cette approche en étudiant les bornes inférieures de la probabilité a posteriori et du facteur de Bayes sur des ensembles de famille de lois a priori. Ces bornes inférieures sont considérées comme réponses bayésiennes les moins favorables pour le test. On considère les deux situations suivantes:

a) Cas des échantillons *i.i.d.*

Casella et Berger (1987) montrent que dans le cas d'un test bilatéral

$$H_0 : \theta \leq \theta_0 \quad \text{contre} \quad H_1 : \theta > \theta_0$$

sur le paramètre de position de densité $f(x/\theta)$, la réponse classique coïncide avec la réponse bayésienne exprimée en terme de la probabilité a posteriori de l'hypothèse H_0 .

En effet, dans plusieurs classes de lois a priori considérées, les auteurs ont montré que la borne inférieure de la probabilité a posteriori de l'hypothèse H_0 est souvent égale ou inférieure à la p-valeur. Contrairement au test bilatéral suivant:

$$H_0 : \theta = \theta_0 \quad \text{contre} \quad H_1 : \theta \neq \theta_0,$$

Berger et Sellke (1987) montrent que la p-valeur est plus petite que la borne inférieure de la probabilité a posteriori ce qui explique qu'il ya un désaccord entre la réponse classique et la réponse bayésienne au test.

b) Cas des échantillons *non i.i.d.*

Berger et Yang(1994) comparent la réponse bayésienne à la réponse classique pour le test de la racine unitaire dans le modèle AR(1) suivant

$$X_t = \rho X_{t-1} + \epsilon_t,$$

où les ϵ_t sont des variables aléatoires *i.i.d* de loi $N(0, \sigma^2)$.

Le test d'hypothèse consiste à tester les trois d'hypothèses suivantes deux à deux.

$$H_0 : \rho = 1 \quad \text{contre} \quad H_1 : \rho < 1 \quad \text{contre} \quad H_2 : \rho > 1$$

Les auteurs considèrent les lois a priori symétriques de référence

$$g_1 = 1/[\pi\sqrt{1-\rho^2}] \quad \text{et} \quad g_2 = 1/[\pi|\rho|\sqrt{\rho^2-1}]$$

ainsi que les classes de lois a priori objectives suivantes:

$$G_1 = \{\text{uniforme sur } (1-r, 1), \quad r > 0\}, \text{ pour } 0 < \rho < 1.$$

$$G_2 = \{\text{uniforme sur } (1, 1+r), \quad r > 0\}, \text{ pour } \rho > 1.$$

Ils développent les expressions des bornes inférieures des facteurs de Bayes suivants: B_{01} , B_{02} , et B_{12} .

Les résultats de l'étude montrent que dans le cas du test unilatéral de type

$$H_0 : \rho = 1 \quad \text{contre} \quad H_1 : \rho < 1,$$

ou

$$H_0 : \rho = 1 \quad \text{contre} \quad H_2 : \rho > 1$$

La borne inférieure du facteur de Bayes B_{01} ou B_{02} est plus grande que la p-valeur. Cette situation correspond au résultat de Berger et Sellke (1987).

Par contre dans le cas du test bilatéral

$$H_1 : \rho < 1 \quad \text{contre} \quad H_2 : \rho > 1$$

il a été constaté que la borne inférieure du facteur de Bayes B_{12} est plus grande que la p-valeur.

Ceci permet de faire la conclusion suivante: La généralisation du résultat de Casella et Berger (1987) n'est pas valable dans le cas des échantillons non *i.i.d.*

Critique faite sur la réponse la moins favorable.

La réponse bayésienne la moins favorable à l'hypothèse nulle ne permet pas de conclure en défaveur de l'hypothèse nulle.

7 Le paradoxe de Jeffreys-Lindley-Bartlett.

Les tests bayésiens d'une hypothèse nulle simple souffrent d'un paradoxe découvert par Jeffreys (1939), Lindley (1957) et Bartlett (1957). Ce paradoxe semble indiquer que les lois a priori (propres et impropres) sont non appropriées pour un test d'hypothèse nulle simple dans certains cas. A titre d'exemple, le test de la moyenne d'une loi normale.

Exemple 7.1:

Lindley (1957), se donne une variable aléatoire $X \sim N(\theta, 1)$, où θ est un paramètre inconnu à tester et considère le test d'hypothèses

$$H_0 : \theta = \theta_0 \quad \text{contre} \quad H_1 : \theta \neq \theta_0.$$

L'auteur démontre le résultat suivant

Proposition 7.1 (Lindley, 1957)

Si la loi a priori est la loi uniforme sur l'intervalle $[-\alpha, \alpha]$ de densité

$$\pi(\theta) = \frac{1}{2\alpha}$$

alors

i) la probabilité a posteriori de H_0 s'écrit:

$$\pi(\theta = \theta_0/X) = [1 + \frac{1 - P(H_0)}{P(H_0)} \times \frac{1}{B_{01}}]^{-1}$$

ii) Le facteur de Bayes en faveur de H_0 s'exprime par

$$B_{01} = \frac{L(X/\theta = \theta_0)}{\int_{-\alpha}^{+\alpha} \frac{1}{2\alpha} L(X/\theta) d\theta}$$

iii) $\lim_{\alpha \rightarrow \infty} B_{01} = \infty$.

Ce dernier résultat se traduit par l'acceptation de l'hypothèse nulle H_0 systématiquement en faisant élargir le support de la loi a priori.

Bartlett (1957) reprend le même test en considérant comme loi a priori $\pi(\theta) \sim N(0, K)$ et retombe sur le paradoxe de Jeffreys (1939) (i.e: lorsque $K \longrightarrow \infty$, $B_{01} \longrightarrow \infty$).
Jeffreys (1939) utilise la loi a priori de Cauchy et découvre le même paradoxe.

Exemple 7.2.

Ce paradoxe a été retrouvé également dans le problème du test sur les paramètres d'un modèle de régression simple. Conigliani et spezzaferri (2007) montrent que l'utilisation de loi a priori impropre fait apparaître le paradoxe de J-L-B. Celle-ci rend le calcul du facteur de Bayes non approprié.

Résolution du paradoxe de J-L-B.

Le fait que le même paradoxe surgit avec les trois lois a priori dans l'exemple 7.1, conduit Jeffrey (2008) à penser que la loi a priori peut ne pas être la vraie source du problème. En effet, Jeffrey (2008) a présenté des résolutions à certains paradoxes notamment celui présenté en exemple 7.1. Pour lui ce paradoxe vient de la spécification imprécise des hypothèses. Ainsi, il propose de tester l'hypothèse:

$$H_0 : |\theta - \theta_0| \leq \epsilon \quad \text{contre toutes les hypothèses alternatives} \quad H_i : |\theta - \theta_i| \leq \epsilon, i = 1, 2, \dots, \infty$$

et de calculer le facteur de Bayes en faveur de l'hypothèse H_0 contre H_i lorsque $\epsilon \longrightarrow 0$.
En effet,

$$B_{0i} = \frac{\pi(|\theta - \theta_0| < \epsilon/X)}{\pi(|\theta - \theta_i| < \epsilon/X)} \longrightarrow \frac{\pi(\theta_0/X)}{\pi(\theta_i/X)} \quad \text{lorsque} \quad \epsilon \longrightarrow 0.$$

La borne inférieure du facteur de Bayes sera égale à

$$\underline{B}_{0i} = \inf_{\theta} B_{0i} = \frac{\pi(\theta_0/X)}{\sup_{\theta} \pi(\theta_i/X)}.$$

Jeffrey(2008) établit des conditions suffisantes de l'apparition de ce paradoxe qui sont reliées aux hypothèses du test et de la densité a posteriori du paramètre à tester.

Chapitre 2:

Test bayésien de la racine unitaire dans le modèle autorégressif d'ordre 1

1 Introduction.

Dans ce chapitre, nous nous intéressons au test bayésien de la racine unitaire dans le modèle AR(1) stationnaire. Certains auteurs ont développé des tests bayésiens dans le même modèle dont le premier fut Sims (1988). Dejong et Whiteman (1991) apportèrent les premières applications empiriques d'où il ressortait que dans l'optique bayésienne, les séries macro-économiques américaines étaient en majorité stationnaires, contrairement aux résultats trouvés par Nelson et Plosser (1982) qui utilisaient une approche classique. Cette divergence des résultats a pu être expliquée par le fait que les tests classiques de la racine unitaire sont souvent très peu puissants. Cependant, après l'apparition de l'article de Phillips (1991), la robustesse des tests bayésiens fut déclarée dépendre du choix approprié de la loi a priori. En effet, la loi uniforme utilisée sur le paramètre du modèle favorisait trop souvent l'hypothèse de la racine unitaire.

Schotman et Van Dijk (1991, 1993) ont développé un test basé sur des posterior odds. Ils trouvent des résultats finaux similaires à ceux de Phillips sur les données de Nelson et Plosser. Mais leurs résultats restent fragiles car les posterior odds dépendent de l'a priori qui est choisie.

Lubrano (1995) considère également le test bayésien de la racine unitaire et insiste sur le fait que les conditions initiales du processus doivent être prises en compte et modélisées. Il utilise la loi Beta symétrique comme loi a priori du paramètre qui a conduit également à favoriser beaucoup moins souvent l'hypothèse de la racine unitaire comparant à l'a priori uniforme.

Berger et Yang (1994) proposent des lois a priori non informatives pour le test de la racine unitaire pour le modèle AR(1) sans intercept appelées lois de références. Au moyen du facteur de Bayes, les auteurs établissent le lien avec l'approche classique en vérifiant la validité des résultats de Berger et Sellke (1987) dans le cas du test unilatéral (voir chapitre 1).

Marriott et Newbold (1998) abordent le problème du test bayésien de la racine unitaire comme un problème bayésien de sélection entre deux modèles: le modèle de marche aléatoire et le modèle stationnaire. Ils développent alors un test bayésien asymptotiquement robuste basé sur loi Beta non symétrique qui assigne la plus grande masse de probabilité au point $\rho = 1$. Suivant leurs démarche, nous proposons une nouvelle loi a priori issue de la loi exponentielle. Le test construit est asymptotiquement robuste et demeure robuste pour le cas des petites tailles d'échantillons.

Par ailleurs, dans ce chapitre, nous comparons les différents tests bayésiens aux tests classiques (test de Dickey-Fuller et test du maximum de vraisemblance) au moyen de la puissance en incluant le cas des petits échantillons.

2 Position du problème.

Marriott et Newbold (1998) abordent le test bayésien de la racine unitaire suivant:

$$H_0 : \rho = 1 \quad \text{contre} \quad H_1 : \rho < 1 \quad (1)$$

dans le modèle AR(1) avec intercept

$$X_t - \mu = \rho(X_{t-1} - \mu) + \epsilon_t \quad (2)$$

où les ϵ_t sont *i.i.d.* $N(0, \sigma^2)$ et μ est un paramètre inconnu, comme un problème bayésien de sélection entre les deux modèles suivants:

Le modèle de marche aléatoire

$$X_t = X_{t-1} + \epsilon_t \quad ,$$

et le modèle stationnaire

$$X_t - \mu = \rho(X_{t-1} - \mu) + \epsilon_t, \quad |\rho| < 1$$

L'état initial X_0 est supposé égal à zéro.

Le fait que la moyenne μ du processus est inconnue complique la procédure du test bayésien. En effet, le paramètre μ apparaît simplement dans le modèle stationnaire. Dans le modèle

de marche aléatoire, le paramètre n'existe pas. Ainsi, il ne peut pas être considéré comme un vrai paramètre de nuisance comme le cas du paramètre σ .

Schotman et Van Dijk (1991) utilisent la loi uniforme non informative comme loi a priori pour μ sur un intervalle fini de densité:

$$f(\mu) = 1/M, \quad -M/2 < \mu < M/2$$

en supposant μ et ρ deux paramètres indépendants.

Une conséquence de la non identification du paramètre μ dans le modèle de marche aléatoire est la divergence de l'odds ratio a posteriori lorsque $M \rightarrow \infty$.

Afin de pallier à ce problème, Marriott et Newbold (1998) proposent d'éliminer le paramètre μ en considérant l'échantillon (W_1, \dots, W_n) de moyenne nulle à la place de l'échantillon (X_1, \dots, X_n) avec

$$W_t = X_t - X_{t-1}, \quad \forall t = 1, \dots, n.$$

Les auteurs transforment alors le problème de test à une comparaison entre les deux modèles suivants par l'approche bayésienne

$$M_1 : \quad W_t = \epsilon_t$$

$$M_2 : \quad W_t - \rho W_{t-1} = \epsilon_t - \epsilon_{t-1}.$$

3 Fonctions de vraisemblance des modèles.

Les expressions des fonctions de vraisemblance des modèles M_i , $i=1,2$ ont été obtenues par Newbold (1974). Soit l'échantillon,

$$W = (W_1, W_2, \dots, W_n).$$

Dans le modèle M_1 , la fonction de vraisemblance s'écrit en terme de W_t comme suit

$$P(W/\sigma, M_1) = \left(\frac{1}{2\pi\sigma^2}\right)^{n/2} \exp\left\{-\frac{1}{2\sigma^2} \sum_{i=1}^n W_t^2\right\}.$$

et dans le modèle M_2 ,

$$P(W/\rho, \sigma, M_2) = \left(\frac{1}{2\pi\sigma^2}\right)^{n/2} A^{\frac{1}{2}} \exp\left\{-\frac{1}{2\sigma^2} \sum_{i=1}^{n+2} u_t^2\right\}.$$

où

$$A = [1 + n(1 - \rho)(1 + \rho)^{-1}]^{-1},$$

$$u_1 = -AC,$$

$$u_2 = \rho(1 - \rho^2)^{-\frac{1}{2}}AC,$$

$$u_3 = W_1 - (1 + \rho)^{-1}AC,$$

$$u_t = W_{t-2} + (1 - \rho) \sum_{j=1}^{t-3} W_j (1 + \rho)^{-1}AC; \quad t = 4, \dots, n + 2.$$

avec

$$C = (1 - \rho) \sum_{t=1}^n W_t + (1 - \rho)^2 \sum_{i=1}^n (n - t)W_t.$$

4 Etude a posteriori.

4.1 Les densités jointes intégrées des modèles.

Marriott et Newbold (1998) utilisent la loi a priori non informative pour σ . Ils posent

$$P(\rho, \sigma) = \frac{1}{\sigma} p(\rho).$$

où $p(\rho)$ est la loi a priori assignée au paramètre ρ . Les auteurs développent l'expression des densités jointes intégrées $P(W/M_i)$, pour $i=1,2$ comme suit:

En intégrant $P(W/\sigma, M_1)$ par rapport à σ , ils obtiennent

$$P(W/M_1) = (2\pi)^{-\frac{n}{2}} \cdot \Gamma\left(\frac{n}{2}\right) \cdot 2^{\frac{n-2}{2}} \cdot \left[\sum_{t=1}^n W_t^2\right]^{-\frac{n}{2}}.$$

et

$$P(W/M_2) = \int_{-1}^1 P(\rho, W/M_2) \, d\rho$$

avec

$$P(\rho, W/M_2) = (2\pi)^{-\frac{n}{2}} \cdot \Gamma\left(\frac{n}{2}\right) \cdot 2^{\frac{n-2}{2}} \cdot A^{\frac{1}{2}} \cdot \left[\sum_{t=1}^{n+1} u_t^2\right]^{-\frac{n}{2}} \cdot p(\rho)$$

obtenue par intégration de la densité jointe de (ρ, σ, W) par rapport à σ qui est donnée comme suit

$$P(\rho, \sigma, W/M_2) = \frac{1}{\sigma^{n+1}} \left(\frac{1}{2\pi}\right)^{\frac{n}{2}} A^{\frac{1}{2}} \exp\left\{-\frac{1}{2\sigma^2} \sum_{t=1}^{n+2} u_t^2\right\} p(\rho)$$

4.2 Les probabilités a posteriori des modèles.

Le calcul des probabilités a posteriori des modèles $P(M_i/W)$, $i = 1, 2$ est donné par Marriott et Newbold (1998). Pour ce faire, les auteurs utilisent le théorème de Bayes suivant:

$$P(M_i/W) = \frac{P(M_i)P(W/M_i)}{\sum_{i=1}^2 P(W/M_i)P(M_i)}, \quad i = 1, 2$$

en supposant a priori qu'aucun des deux modèles n'est favorisé. Ce qui est traduit par l'égalité suivante:

$$P(M_1) = P(M_2) = \frac{1}{2}.$$

Le calcul des probabilités a posteriori des modèles a conduit aux expressions (3) et (4).

$$P(M_1/W) = \frac{1}{1 + K} \quad (3)$$

et

$$P(M_2/W) = \frac{K}{1 + K} \quad (4)$$

avec

$$K = \frac{\int_{-1}^1 A^{\frac{1}{2}} [\sum_{t=1}^{n+2} u_t]^{-\frac{n}{2}} p(\rho) d\rho}{[\sum_{t=1}^n W_t^2]^{-\frac{n}{2}}}$$

La valeur $K' = \frac{1}{K}$ représente le quotient d'enjeux a posteriori (posterior odds ratio).

5 Règle de sélection des modèles.

La règle de décision la plus simple considérée est qu'on accepte le modèle M_i si la probabilité $P(M_i/W) > 0.5$ pour $i = 1, 2$ pour un échantillon $W = (W_1, \dots, W_n)$ donné.

Pour une série d'échantillons donnée, nous calculons la proportion d'échantillons pour lesquels $P(M_i/W) > 0.5$ pour les différentes lois a priori. La loi a priori la plus appropriée sera alors celle qui correspond à la plus grande proportion d'échantillons.

6 Les lois a priori du test.

Dans la littérature, plusieurs lois a priori pour le paramètre du modèle (2) sont proposées.

• **Loi a priori de Schotman et Van Dijk .**

Schotman et Van Dijk (1991a) utilisent la loi uniforme comme loi a priori pour ρ de densité donnée par:

$$\begin{cases} p(\rho = 1) = v \\ p(\rho, \rho \in S) = \frac{1}{1-a}. \end{cases}$$

avec $S = \{\rho/ -1 < a \leq \rho < 1\}$ et $\rho \in S \cup \{1\}$.

Les auteurs ont obtenus l'expression de l'odds ratio a posteriori correspondant au test

$$H_0 : \rho = 1 \quad \text{contre} \quad H_1 : \rho \in S.$$

Ils établissent une étude empirique sur des données représentant des taux de change.

• **Loi a priori de Berger et Yang .**

Berger et Yang (1994) proposent pour le paramètre du modèle (2) sans terme constant la loi a priori non informative de référence qui s'écrit :

$$\pi_{SR}(\rho) = \begin{cases} 1/(2\pi\sqrt{1-\rho^2}) & \text{si } |\rho| < 1 \\ 1/(2\pi |\rho| \sqrt{\rho^2-1}) & \text{si } |\rho| > 1 \end{cases}$$

ainsi que la loi a priori de Jeffrey qui s'écrit dans ce cas comme suit

$$\pi_J(\rho) = \frac{n}{1-\rho^2} + \frac{1-\rho^{2n}}{1-\rho^2} \left\{ E\left[\frac{X_0^2}{\sigma^2}\right] - \frac{1}{1-\rho^2} \right\}^{1/2}$$

pour le test d'hypothèses multiples:

$$H_0 : \rho = 1, \quad H_1 : \rho < 1 \quad \text{et} \quad H_2 : \rho > 1$$

Les auteurs procèdent à une étude comparative de la réponse bayésienne (facteur de Bayes) et de la réponse classique (p-valeur) au problème de test. Ainsi, ils comparent leur résultats avec ceux de Berger et Sellke (1987) et Casella et Berger (1987) (voir section 6 du chapitre1).

• **Loi a priori de Lubrano .**

Lubrano (1995) considère comme loi a priori pour ρ , la loi Beta symétrique sur l'intervalle $[-\sqrt{1+v}, \sqrt{1+v}]$ qui s'écrit comme suit

$$p(\rho/v) \propto (\sqrt{1+v} + \rho)^{p-1} (\sqrt{1+v} - \rho)^{q-1}, \quad v > 0$$

L'auteur fait donc une comparaison de ces résultats avec ceux obtenus par Schotman et Van Dijk, en montrant que l'état initial X_0 joue un rôle important sur les conclusions du test de la racine unitaire.

• **La loi a priori de Marriott et Newbold.**

Les auteurs considèrent la loi Beta de densité

$$p(\rho) = \frac{\Gamma(\alpha + \beta)}{\Gamma(\alpha)\Gamma(\beta)2^{\alpha+\beta-1}} (1 + \rho)^{\alpha-1} (1 - \rho)^{\beta-1}, \quad |\rho| < 1$$

pour $(\beta = 0.5, \alpha = 5)$ et pour $(\beta = 0.5, \alpha_2 = 50)$. On note ces distributions par l'a priori Beta1 et l'a priori Beta2 respectivement.

Les auteurs abordent une étude objective par le moyen des probabilités a posteriori des modèles $M_i, i = 1, 2$. Le calcul de ces probabilités a permis aux auteurs de conclure que lorsque le processus généré est une marche aléatoire, alors la loi a priori uniforme est la plus appropriée pour le test. Dans le cas contraire, c'est-à-dire lorsque le processus généré est stationnaire, les lois a priori Beta2 et Beta1 données dans cet ordre sont les plus appropriées pour le test.

• **Loi a priori de Conigliani et Spezzaferri.**

Conigliani et Spezzaferri (2007) ont proposé la distribution triangulaire pour le paramètre ρ de densité

$$p(\rho) = \frac{1 + \rho}{2}, \quad |\rho| < 1$$

pour le test (1).

Les auteurs abordent le test bayésien de la racine unitaire dans le cas d'un AR(1) et dans le cas d'un ARMA(1,1). Ainsi, ils généralisent l'étude établie par Marriott et Newbold (1998) dans le cadre bayésien au moyen du facteur de Bayes fractionné (O'Hagan, 1995). Par ailleurs, ils étalent l'étude à une comparaison des tests bayésiens et les tests classiques de unit root dans les modèles AR(1) et ARMA(1,1) au moyen de la puissance.

• **La loi a priori proposée.**

Nous proposons comme loi a priori pour ρ , une loi issue de la loi exponentielle de paramètre λ qui est de densité

$$p(\rho) = \frac{\lambda}{1 - \exp\{-2\lambda\}} \exp\{-\lambda(1 - \rho)\}, \quad |\rho| < 1.$$

qu'on désigne par loi exponentielle modifiée de type 1 notée T1Mexp. Nous avons obtenu cette loi par translation de la loi exponentielle de paramètre λ tronquée par la suite entre -1 et +1.

Le principe de construction de cette loi est le même considéré par Marriott et Newbold (1998) qui consiste à proposer une loi a priori qui attribue une grande masse de probabilité au point $\rho = 1$.

Ici, nous considérons deux distributions exponentielles modifiées de type 1 notées T1Mexp1 (cas $\lambda = 5$) et T1Mexp2 (cas $\lambda = 50$).

Nous donnons dans la table1, les valeurs de la moyenne et de la variance des différentes lois a priori.

Table1. Moyenne et variance des lois a priori de ρ .

Lois a priori pour ρ	moyenne	variance
Beta1	0.8181	0.0508
T1Mexp1	0.8009	0.0398
Beta2	0.9800	0.00076
T1Mexp2	0.9800	0.0004

Les figures suivantes représentent les allures des densités des lois a priori.

L'allure des densités des lois a priori de ρ .

7 Etude comparative.

Dans ce paragraphe, nous comparons les tests bayésiens au moyen des probabilités a posteriori. Nous comparons ensuite les tests bayésiens avec certains tests classiques au moyen de la puissance.

7.1 Les probabilités a posteriori des modèles.

Nous simulons la proportion d'échantillons pour lesquels $P(M_1/W) > 0.5$ qui représente la proportion pour laquelle le modèle M_1 est choisi. Les résultats sont basés sur 1000 répétitions. Ceux qui correspondent aux petits échantillons sont donnés dans la table 2. La table 3 est réservée au cas des échantillons de grandes tailles.

Table 2. Proportion d'échantillons pour lesquels $P(M_1/W) > 0.5$.

n	ρ	Lois a priori pour ρ			
		Uniforme	Triangulaire	Beta1	T1Mexp1
10	1.00	<u>0.8720</u>	<u>0.7890</u>	<u>0.6850</u>	<u>0.5460</u>
	0.95	<u>0.8360</u>	<u>0.7360</u>	<u>0.6110</u>	0.4630
	0.90	<u>0.7910</u>	<u>0.6850</u>	<u>0.5490</u>	0.3750
	0.85	<u>0.7560</u>	<u>0.6330</u>	0.4840	0.3020
	0.80	<u>0.7200</u>	<u>0.5880</u>	0.4180	0.2320
	0.50	0.4360	0.2820	0.1370	0.0430
	0.30	0.2550	0.1450	0.0440	0.0150
	0	0.0890	0.0310	0.0120	0.0060
20	1.00	<u>0.9420</u>	<u>0.8230</u>	<u>0.7370</u>	<u>0.6530</u>
	0.95	<u>0.9050</u>	<u>0.7220</u>	<u>0.6090</u>	0.4870
	0.90	<u>0.8730</u>	<u>0.6110</u>	0.4870	0.3490
	0.85	<u>0.8190</u>	<u>0.5100</u>	0.3690	0.2530
	0.80	<u>0.7520</u>	0.4020	0.2760	0.1710
	0.50	0.2800	0.0500	0.0210	0.0070
	0.30	0.0840	0.0070	0.0040	0.0010
	0	0.0070	0.0010	0.0010	0.0000
25	1.00	<u>0.9240</u>	<u>0.8090</u>	<u>0.7590</u>	<u>0.6680</u>
	0.95	<u>0.8790</u>	<u>0.6930</u>	<u>0.6190</u>	0.4950
	0.90	<u>0.8160</u>	<u>0.5470</u>	0.4700	0.3450
	0.85	<u>0.7270</u>	0.4350	0.3470	0.2310
	0.80	<u>0.6270</u>	0.3110	0.2300	0.1560
	0.50	0.1180	0.0160	0.0060	0.0040
	0.30	0.0160	0.0020	0.0010	0.0000
	0	0.0010	0.0000	0.0000	0.0000

- Lorsque le processus généré est une marche aléatoire, nous constatons que la loi a priori uniforme est la plus appropriée que les autres lois a priori car celle-ci donne lieu à une grande proportion d'échantillon (exemple pour $n=10$, la valeur est 0.8720).
- Dans le cas des petits échantillons, nous retenons la loi a priori T1Mexp1 comme étant la plus appropriée que les autres lois a priori lorsque le processus généré est stationnaire.

Remarque:

Les loi a priori Beta2 et T1Mexp2 ne sont pas incluses dans l'étude des petits échantillons pour des raisons de difficultés rencontrées dans la programmation.

Table 3. Proportion d'échantillons pour lesquels $P(M_1/W) > 0.5$.

n	ρ	Lois a priori pour ρ					
		Uniforme	Triangulaire	Beta1	T1Mexp1	Beta2	T1Mexp2
100	1.00	<u>0.9780</u>	<u>0.9720</u>	<u>0.9140</u>	<u>0.9090</u>	<u>0.7740</u>	<u>0.5960</u>
	0.95	<u>0.8690</u>	<u>0.8430</u>	<u>0.5610</u>	<u>0.5500</u>	0.2370	0.0480
	0.90	<u>0.6220</u>	<u>0.5650</u>	0.1870	0.1770	0.0340	0.0050
	0.85	0.2630	0.2170	0.0290	0.0240	0.0030	0.0000
	0.80	0.0710	0.0450	0.0040	0.0030	0.0000	0.0000
	0.50	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
150	1.00	<u>0.9970</u>	<u>0.9950</u>	<u>0.9100</u>	<u>0.9310</u>	<u>0.7670</u>	<u>0.6640</u>
	0.95	<u>0.9310</u>	<u>0.9170</u>	0.3170	0.3760	0.0850	0.0360
	0.90	<u>0.6390</u>	<u>0.5910</u>	0.0250	0.0340	0.0030	0.0000
	0.85	0.2540	0.2120	0.0010	0.0010	0.0000	0.0000
	0.80	0.0310	0.0220	0.0000	0.0000	0.0000	0.0000
	0.50	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
200	1.00	<u>0.9980</u>	<u>0.9980</u>	<u>0.8790</u>	<u>0.8930</u>	<u>0.7880</u>	<u>0.7260</u>
	0.95	<u>0.9570</u>	<u>0.9470</u>	0.1150	0.1310	0.0300	0.0080
	0.90	<u>0.6710</u>	<u>0.6400</u>	0.0000	0.0000	0.0000	0.0000
	0.85	0.1700	0.1390	0.0000	0.0000	0.0000	0.0000
	0.80	0.0030	0.0030	0.0000	0.0000	0.0000	0.0000
	0.50	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000

- Dans le cas asymptotique, lorsque le processus généré est stationnaire, les lois a priori Beta2 et T1Mexp2 sont les plus performantes pour le test que les autres lois (voir table 3).
- Pour les grands échantillons, la loi a priori uniforme demeure la plus appropriée lorsque le processus généré est une marche aléatoire (exemple, $n=100$, la valeur est 0.9780).
- Dans les deux cas (asymptotique ou non), nous enregistrons un inconvénient dans l'utilisation des lois a priori uniforme et triangulaire. Celles-ci favorisent si souvent l'hypothèse de la racine unitaire lorsque le processus généré est stationnaire (exemple: $\rho = 0.9 \forall n$, voir table 2 et table 3).
- Par contre, plus n augmente et lorsque le processus généré est stationnaire, les lois a priori Beta1 et T1Mexp1 favorisent moins souvent l'hypothèse de unit root comparées à la loi uniforme et la loi triangulaire (exemple, pour $n=100$ et $\rho = 0.90$, la proportion pour Beta1 est 0.1870, pour la loi uniforme est 0.6220. Pour $n=200$, et $\rho = 0.9$ la proportion pour Beta1 est 0, pour la loi uniforme est 0.6710).

Au moyen de la fonction de répartition empirique de la variable $P(M_1/W)$, nous pouvons retrouver les résultats précédents.

Dans ce qui suit, nous donnons les allures des fonctions de répartition de $P(M_1/W)$ pour les différentes lois a priori considérées dans le cas des petits et grands échantillons.

Cas des petits échantillons

Fonctions de répartitions empiriques de $P(M_1 = W)$

Fonctions de répartitions empiriques de $P(M_1 = W)$

Nous pouvons remarquer dans le cas des petits échantillons que la fonction de répartition de $P(M_1 = W)$ pour la loi a priori T1Mexp1 montre une croissance élevée autour de la médiane et atteint la valeur 1 plus vite que les autres lois a priori lorsque le processus généré est stationnaire.

Cas des grands échantillons($n=100$).

Fonctions de répartitions empiriques de $P(M_1 = W)$

Fonctions de répartitions empiriques de $P(M_1 = W)$

7.2 La puissance des tests.

Le seuil du test bayésien de type (1) est défini par

$$\alpha = P_{H_0}(M_2/W) = 1 - P_{H_0}(M_1/W),$$

Sa puissance $\pi(\cdot)$ est

$$\pi(\rho) = 1 - \beta(\rho) = P_{H_1}(M_2/W) = 1 - P_{H_1}(M_1/W),$$

Conigliani et Spezzaferri (2007) ont étudié également la sensibilité du test (1) par rapport au choix de la loi a priori par l'attribution d'une nouvelle loi a priori, la distribution triangulaire. Les auteurs ont comparé les tests bayésiens aux tests classiques; test de Dickey-Fuller (D-F) et du maximum de vraisemblance (MLE) par rapport à la puissance dans le cas des grands échantillons.

Dans cette partie, nous élargissons l'étude en incluant notre test bayésien basé sur la loi a priori T1Mexp1 et T1Mexp2. L'étude des petits échantillons étant aussi considérée dans cette partie.

La statistique de test de Dickey-Fuller s'écrit:

$$\hat{\tau}_{OLS} = [\hat{\sigma}_{OLS}^{-2} \sum_{t=2}^n (X_{t-1} - \bar{X}_{(-1)})^2]^{1/2} (\hat{\rho}_{OLS} - 1)$$

où

$\hat{\rho}_{OLS}$ et $\hat{\sigma}_{OLS}^2$ sont les estimateurs des moindres carrés de ρ et σ^2 respectivement.

avec

$$\hat{\rho}_{OLS} = [\sum_{t=2}^n (X_{t-1} - \bar{X}_{(-1)})^2]^{-1} \sum_{t=2}^n (X_{t-1} - \bar{X}_{(-1)}) X_t$$

et

$$\hat{\sigma}_{OLS}^2 = (n-3)^{-1} \sum_{t=2}^n [X_t - \bar{X}_{(0)} - \hat{\rho}_{OLS}(X_{t-1} - \bar{X}_{(-1)})]^2$$

où $\bar{X}_{(0)} = (n-1)^{-1} \sum_{t=2}^n X_t$, et $\bar{X}_{(-1)} = (n-1)^{-1} \sum_{t=2}^n X_{t-1}$.

La statistique de test du maximum de vraisemblance:

$$\hat{\tau}_{MV} = [\hat{\sigma}_{MV}^{-2} \sum_{t=2}^n (X_{t-1} - \hat{\mu}_{MV})^2]^{1/2} (\hat{\rho}_{MV} - 1)$$

où $\hat{\rho}_{MV}$, $\hat{\mu}_{MV}$ et $\hat{\sigma}_{MV}^2$ sont les estimateurs du maximum de vraisemblance des paramètres ρ , μ et σ^2 respectivement.

avec

$$\hat{\sigma}_{MV}^2 = n^{-1} (X_1 - \hat{\mu}_{MV})^2 + n^{-1} \sum_{t=2}^n [X_t - \hat{\mu}_{MV}(1 - \hat{\rho}_{MV}) - \hat{\rho}_{MV} X_{t-1}]^2$$

$$\hat{\mu}_{MV} = \frac{X_1 + (1 - \hat{\rho}_{MV}) \sum_{t=2}^n (X_t - \hat{\rho}_{MV} X_{t-1})}{1 + (n-1)(1 - \hat{\rho}_{MV})^2}$$

et

$$\hat{\rho}_{MV} = \frac{\sum_{t=2}^n (X_t - \hat{\mu}_{MV})(X_{t-1} - \hat{\mu}_{MV})}{(X_{t-1} - \hat{\mu}_{MV})}$$

La distribution limite des statistiques de test sont données dans les propositions suivantes:

Proposition 7.2 (Pantula et al, 1994).

La statistique du test des moindres carrés $\hat{\tau}_{OLS}$ converge en distribution vers

$$(G - H^2)^{-1/2}(\xi - TH).$$

où $\xi = 0.5(T^2 - 1)$, $T = W(1)$, $G = \int_0^1 W^2(t)dt$ $H = \int_0^1 W(t)dt$.
avec $W(t)$ est le mouvement Brownien sur $[0, 1]$.

Proposition 7.3 (Pantula et al, 1994).

La statistique du maximum de vraisemblance $\hat{\tau}_{MV}$ converge en distribution vers

$$G^{-1/2}\xi$$

où $\xi = 0.5(T^2 - 1)$, $T = W(1)$, $G = \int_0^1 W^2(t)dt$ $H = \int_0^1 W(t)dt$.
et $W(t)$ est le mouvement Brownien sur $[0, 1]$.

Les tableaux 5 et 6 donnent les valeurs du seuil et de la puissance des différents tests bayésiens et classiques. Les valeurs de la puissance des différents tests classiques (test de Dickey-Fuller et test du maximum de vraisemblance) sont tirées de la table de Pantula et al (1994).

	ρ	Lois a priori pour ρ				test D-F	test MLE
		Uniforme	Triangulaire	Beta1	T1Mexp1		
α	1	0.076	0.191	0.410	0.332	0.05	0.05
$1 - \beta$	0.95	0.121	0.307	0.381	0.505	0.0622	0.085
$1 - \beta$	0.90	0.184	0.453	0.530	0.655	0.0694	0.102
$1 - \beta$	0.85	0.273	0.565	0.653	0.769	0.0846	0.148
$1 - \beta$	0.80	0.373	0.689	0.770	0.844	0.121	0.205

Table 5. Les valeurs de la puissance des différents tests, n=25.

- Comme le montre la table 5, les différents test bayésiens considérés sont plus puissants que les tests classiques (tests de D-F et MLE) pour les petites tailles d'échantillons.
- De la table 5 et 6, on peut dire que le seuil du test bayésien n'étant pas une valeur fixe. Celui-ci dépend de la loi a priori assignée au test. Tandis que pour les tests classiques (test de D-F et test du MLE), le seuil est une valeur fixée au préalable et égale à 0.05.

	ρ	Lois a priori pour ρ						test D-F	test MLE
		Uniforme	Triangulaire	Beta1	T1Mexp1	Beta2	T1Mexp2		
α	1	0.022	0.028	0.086	0.091	0.226	0.404	0.05	0.05
$1 - \beta$	0.95	0.131	0.157	0.439	0.450	0.763	0.952	0.117	0.276
$1 - \beta$	0.90	0.378	0.435	0.813	0.823	0.966	0.995	0.310	0.644
$1 - \beta$	0.85	0.737	0.873	0.971	0.976	0.997	1	0.621	0.905
$1 - \beta$	0.80	0.929	0.955	0.996	0.997	1	1	0.877	0.988

Table 6. Les valeurs de la puissance des différents tests, n=100.

- Dans le cas asymptotique, les tests bayésiens basés sur les lois a priori uniforme et triangulaire semblent être moins puissants que le test classique basé sur la statistique du maximum de vraisemblance (test du MLE).
- La plus grande puissance enregistrée correspond au test basé sur la loi a priori T1Mexp2 pour les grands échantillons et sur la loi a priori T1Mexp1 pour les petits échantillons.
- Pour la même raison évoquée en section 7.1, nous n'avons pas inclu l'étude avec les lois a priori

Beta2 et T1Mexp2 pour le cas des petits échantillons.

Selon le premier critère utilisé qui est la probabilité a posteriori des modèles, nous concluons que lorsque le processus généré est une marche aléatoire, nous retenons la loi uniforme comme étant la plus appropriée dans le cas des grands et petits échantillons.

Dans le cas contraire, c'est-à-dire lorsque le processus généré est stationnaire, la loi a priori T1Mexp2 puis la Beta2 (dans cet ordre) sont plus performantes pour le test dans le cas asymptotique. Dans le cas des petits échantillons, nous enregistrons un comportement meilleur du test basé sur la loi a priori T1Mexp1.

Ceci nous permet de dire qu'à présent, on ne détient pas une loi a priori qui répond favorablement bien aux deux hypothèses du test. En effet, la loi a priori uniforme a tendance à favoriser l'hypothèse de unit root si souvent. Tandis qu'elle est plus appropriée que la loi T1Mexp2 dans le cas asymptotique (T1Mexp1 dans le cas des petits échantillons) lorsque le processus généré est une marche aléatoire.

Chapitre 3:

Test bayésien d'indépendance dans le modèle autorégressif d'ordre 1

1 Introduction.

Le modèle autorégressif d'ordre 1 est l'un des modèles utilisés pour modéliser la dépendance dans une série chronologique. Le paramètre du modèle AR(1) est interprété comme un paramètre d'autocorrélation. L'inférence liée à ce paramètre a attiré l'attention de beaucoup de chercheurs. Les tests d'indépendance permettent justement de détecter la dépendance (présence d'autocorrélation) dans une série d'observations. Durbin et Watson (1950) ont proposé une statistique pour tester l'hypothèse nulle d'absence d'autocorrélation. Basawa et al (1985) ont proposé des statistiques robustifiées pour les modèles autorégressifs en particuliers le modèle AR(1). Andrews (1998) a considéré des tests de corrélation sérielle saisonnière ou non saisonnière. Berkoun et al (2003) ont étudié la robustesse des tests classiques d'indépendance du paramètre autorégressif d'ordre 1 en présence d'un outlier additif dans le modèle.

Dans ce chapitre, nous abordons le test d'indépendance par l'approche bayésienne dans le modèle autorégressif d'ordre 1. Nous proposons différentes lois a priori pour le paramètre du modèle. Ces lois sont construites de sorte à attribuer une plus grande masse de probabilité au point $\rho = 0$.

Dans la première partie du chapitre, nous effectuons une étude comparative des tests proposés au moyen des probabilités a posteriori. La deuxième partie est consacrée à une étude comparative des tests bayésiens d'indépendance et certains tests classiques au moyen de la puissance.

2 Modélisation de l'autocorrélation.

Soit le modèle autorégressif d'ordre 1 suivant:

$$X_t = \rho X_{t-1} + \epsilon_t \quad (1)$$

où les ϵ_t sont des variables aléatoires *i.i.d.* de loi $N(0, \sigma^2)$. L'état initial X_0 est supposé égal à zéro. Le paramètre ρ est nommé paramètre d'autocorrélation et sa valeur doit être dans l'intervalle $[-1, 1]$.

Proposition 7.4 Fuller (1976). *Pour tout $t = 1, 2, \dots$*

- $X_t = \sum_{j=0}^{t-1} \rho^j \epsilon_{t-j}$.
- $\lim_{t \rightarrow \infty} E[X_t] = 0$,
- $\lim_{t \rightarrow \infty} V(X_t) = \frac{\sigma^2}{1 - \rho^2}$
- $\lim_{t \rightarrow \infty} Cov(X_t, X_{t+j}) = \rho^{|j|} \frac{\sigma^2}{(1 - \rho^2)}$.

Comme on peut le voir, une autocorrélation positive entre X_t et X_{t+1} est obtenue lorsque $\rho > 0$, tandis que $\rho < 0$ donne une autocorrélation négative. Il n'y a pas d'autocorrélation lorsque $\rho = 0$.

3 Position du problème.

Nous considérons le test bayésien d'indépendance suivant:

$$H_0 : \rho = 0 \quad \text{contre} \quad H_1 : \rho \neq 0$$

dans le modèle autorégressif d'ordre 1 stationnaire

$$X_t = \rho X_{t-1} + \epsilon_t$$

où les ϵ_t sont des variables aléatoires *i.i.d.* de loi $N(0, \sigma^2)$ avec $|\rho| < 1$. L'état initial X_0 est supposé égal à zéro.

Nous nous proposons d'aborder ce problème comme un problème bayésien de sélection entre les modèles suivants:

$$\begin{cases} M_1 : & X_t = \epsilon_t \\ M_2 : & X_t = \rho X_{t-1} + \epsilon_t, \quad |\rho| < 1, \quad \rho \neq 0 \end{cases}.$$

4 Fonctions de vraisemblance des modèles.

Soit l'échantillon,

$$X = (X_1, \dots, X_n).$$

La fonction de vraisemblance sous le modèle M_1 s'écrit:

$$P(X/\sigma, M_1) = \left(\frac{1}{2\pi\sigma^2}\right)^{n/2} \exp\left\{-\frac{1}{2\sigma^2} \sum_{t=1}^n X_t^2\right\}$$

Pour le modèle M_2 :

$$P(X/\rho, \sigma, M_2) = \left(\frac{1}{2\pi\sigma^2}\right)^{n/2} \exp\left\{-\frac{1}{2\sigma^2} \sum_{t=1}^n (X_t - \rho X_{t-1})^2\right\}, \quad \rho \neq 0$$

5 Etude a posteriori

5.1 Les densités jointes intégrées des modèles.

Nous considérons la loi a priori non informative pour le paramètre σ .

soit

$$P(\rho, \sigma) = \frac{1}{\sigma} p(\rho)$$

où $p(\rho)$ est la distribution a priori assignée au paramètre ρ .

Pour le modèle M_1 :

Nous intégrons $P(X/\sigma, M_1)$ par rapport à σ , on obtient:

$$P(X/M_1) = (2\pi)^{-n/2} \Gamma\left(\frac{n}{2}\right) 2^{\left(\frac{n-1}{2}\right)} \left[\sum_{t=1}^n (X_t^2)\right]^{-n/2}$$

Pour le modèle M_2 :

La densité jointe de (ρ, σ, X) pour le modèle M_2 s'écrit:

$$P(\sigma, \rho, X/M_2) = (2\pi)^{-n/2} \left(\frac{1}{\sigma}\right)^{n+1} \Gamma\left(\frac{n}{2}\right) 2^{\left(\frac{n-1}{2}\right)} \left[\sum_{t=1}^n (X_t - \rho X_{t-1})^2\right]^{-n/2} p(\rho)$$

Nous intégrons par rapport à σ , nous obtenons

$$P(\rho, X/M_2) = (2\pi)^{-n/2} \Gamma\left(\frac{n}{2}\right) 2^{\left(\frac{n-1}{2}\right)} \left[\sum_{t=1}^n (X_t - \rho X_{t-1})^2 \right]^{-n/2} p(\rho)$$

La densité jointe intégrée de (ρ, σ, X) pour le modèle M_2 s'écrit:

$$\begin{aligned} P(X/M_2) &= \int_{-1}^1 P(\rho, X/M_2) 1_{\{\rho \neq 0\}} d\rho \\ &= (2\pi)^{-n/2} \Gamma\left(\frac{n}{2}\right) 2^{\left(\frac{n-1}{2}\right)} \int_{-1}^1 \left[\sum_{t=1}^n (X_t - \rho X_{t-1})^2 \right]^{-n/2} p(\rho) 1_{\{\rho \neq 0\}} d\rho \end{aligned}$$

5.2 Les probabilités a posteriori des modèles.

Pour le calcul des probabilités a posteriori des modèles, nous utilisons le théorème de Bayes suivant:

$$P(M_i/X) = \frac{P(M_i)P(X/M_i)}{\sum_{i=1}^2 P(X/M_i)P(M_i)}, \quad i = 1, 2$$

Nous supposons que $P(M_1) = P(M_2) = \frac{1}{2}$.

La probabilité a posteriori du modèle M_1 est:

$$P(M_1/X) = \frac{1}{1 + K}$$

La probabilité a posteriori du modèle M_2 s'écrit comme suit:

$$P(M_2/X) = \frac{K}{1 + K}$$

où K est le rapport de vraisemblance de M_2 en faveur de M_1 donné par

$$K = \frac{P(X/M_2)}{P(X/M_1)} = \frac{\int_{-1}^1 \left(\sum_{t=1}^n (X_t - \rho X_{t-1})^2 \right)^{-n/2} p(\rho) 1_{\{\rho \neq 0\}} d\rho}{\left(\sum_{t=1}^n X_t^2 \right)^{-n/2}}$$

8 Les lois a priori du test

Nous proposons les différentes lois a priori suivantes:

- **La loi uniforme** sur $(-1, 1)$.

- **La loi triangulaire** sur $(-1, 1)$ de densité

$$p(\rho) = \begin{cases} 1 - |\rho| & \text{si } |\rho| < 1 \\ 0 & \text{sinon} \end{cases}$$

de moyenne nulle et de variance $= 1/6$

- **La loi normale tronquée** sur $(-1, 1)$ notée $N_{Tr}(0, \sigma^2)$ de densité

$$p(\rho) = \begin{cases} \frac{1}{\sigma\sqrt{2\pi}\text{Erf}(\frac{1}{2\sigma})} \exp\{-\frac{1}{2\sigma^2}\rho^2\} & \text{si } |\rho| < 1 \\ 0 & \text{sinon} \end{cases}$$

avec $\text{Erf}(z) = \frac{2}{\sqrt{\pi}} \int_0^z \exp\{-t^2\} dt$.

de moyenne nulle et de variance égale à

$$\sigma_{Tr}^2 = \sigma^2 - \frac{2\sigma \exp\{-\frac{1}{2\sigma^2}\}}{\sqrt{2\pi} \text{Erf}(\frac{1}{\sqrt{2}\sigma})}.$$

Nous considérons la distribution normale tronquée avec $\sigma = 0.05$ que nous notons N_{Tr} .

- **La loi exponentielle modifiée de type 2.**

C'est la loi qu'on note T2Mexp de densité

$$p(\rho) = \begin{cases} \frac{\lambda}{2(1 - \exp\{-\lambda\})} \exp\{-\lambda |\rho|\} & \text{si } |\rho| < 1 \\ 0 & \text{sinon} \end{cases}$$

de moyenne nulle et de variance

$$\sigma_{Exp}^2 = \frac{2 - (2 + 2\lambda + \lambda^2) \exp\{-\lambda\}}{\lambda^2(1 - \exp\{-\lambda\})}$$

Nous considérons deux lois exponentielles modifiées notées T2Mexp1 (cas $\lambda = 5$) et T2Mexp2 (cas $\lambda = 10$).

La table 1 donne les valeurs de la variance des différentes lois a priori.

Table 1. Moments des lois a priori.

Lois a priori pour ρ	variance
Triangulaire	0.1666
N_{Tr}	0.0025
T2Mexp1	0.0705
T2Mexp2	0.0199

Les figures suivantes représentent l'allure des densités des lois a priori que nous considérons dans cette étude.

7 Etude Comparative .

Dans ce paragraphe, nous comparons les tests bayésiens au moyen des probabilités a posteriori. Par la suite nous établissons une étude comparative des tests bayésiens avec certains tests classiques au moyen de la puissance.

7.1 Les probabilités a posteriori.

Nous simulons la proportion d'échantillons pour lesquels $P(M_1/X) > 0.5$ pour les différentes lois a priori. Les résultats obtenus à 1000 répétitions dans le cas des petits échantillons sont donnés dans la table 2. La table 3 est réservée aux grands échantillons.

On restreint l'interprétation des résultats aux valeurs positives de ρ vu la symétrie des résultats simulés.

Table 2. Proportion d'échantillons pour lesquels $P(M_1/X) > 0.5$.

n	ρ	Lois a priori pour ρ				
		Uniforme	Triangulaire	T2Mexp1	T2Mexp2	N_{Tr}
10	-0.30	0.7280	0.6690	0.6340	0.7250	0.9830
	-0.20	0.8020	0.7410	0.7290	0.8050	0.9920
	-0.10	0.8420	0.7880	0.7660	0.8440	0.9960
	-0.05	0.8550	0.8010	0.7720	0.8620	0.9980
	0	0.8390	0.7860	0.7700	0.8450	0.9960
	0.05	0.8220	0.7890	0.7620	0.8290	0.9960
	0.10	0.8120	0.7670	0.7540	0.8170	0.9940
	0.20	0.7720	0.7120	0.6900	0.7720	0.9900
	0.30	0.6770	0.6180	0.5990	0.6850	0.9820
25	-0.30	0.5700	<u>0.4770</u>	<u>0.4290</u>	<u>0.4320</u>	0.6150
	-0.20	0.7220	0.6490	0.6020	0.6050	0.7690
	-0.15	0.8020	0.7220	0.6790	0.6830	0.8390
	-0.05	0.8940	0.8420	0.7850	0.7870	0.9150
	0	0.9240	0.8660	0.8250	0.8290	0.9450
	0.05	0.9050	0.8460	0.7970	0.8000	0.9320
	0.10	0.8680	0.8040	0.7580	0.7590	0.9030
	0.20	0.7500	0.6750	0.6210	0.6230	0.7880
	0.30	0.5750	<u>0.4920</u>	<u>0.4270</u>	<u>0.4330</u>	0.6310

Dans le cas des petits échantillons, nous constatons que toutes les lois a priori considérées ont tendance à favoriser si souvent le modèle M_1 et aucune loi a priori n'est appropriée lorsque le

processus généré est du modèle M_2 .

Parallèlement, nous enregistrons la loi normale tronquée N_{Tr} comme étant la plus appropriée par rapport aux autres lois lorsque le processus généré est du modèle M_1 (exemple: $n=10$, la proportion est 0.9960).

Table 3. Proportion d'échantillons pour lesquels $P(M_1/X) > 0.5$.

n	ρ	Lois a priori pour ρ				
		Uniforme	Triangulaire	T2Mexp1	T2Mexp2	N_{Tr}
100	-0.25	0.3260	0.2380	0.1700	0.1360	0.1120
	-0.20	<u>0.5240</u>	0.4130	0.3300	0.2740	0.2470
	-0.15	<u>0.7120</u>	<u>0.6290</u>	<u>0.5160</u>	0.4500	0.4170
	-0.10	<u>0.8530</u>	<u>0.7890</u>	<u>0.7100</u>	<u>0.6540</u>	<u>0.6270</u>
	-0.05	<u>0.9250</u>	<u>0.8790</u>	<u>0.8310</u>	<u>0.7900</u>	<u>0.7600</u>
	0	<u>0.9610</u>	<u>0.9320</u>	<u>0.8780</u>	<u>0.8360</u>	<u>0.7990</u>
	0.05	<u>0.9330</u>	<u>0.8820</u>	<u>0.8180</u>	<u>0.7780</u>	<u>0.7580</u>
	0.10	<u>0.8390</u>	<u>0.7830</u>	<u>0.7000</u>	<u>0.6420</u>	<u>0.5960</u>
	0.15	<u>0.7190</u>	<u>0.6120</u>	<u>0.5080</u>	0.4420	0.4090
	0.20	<u>0.5070</u>	0.4030	0.3190	0.2720	0.2370
	0.25	0.3180	0.2320	0.1560	0.1280	0.1140
200	-0.20	0.2530	0.1830	0.1160	0.0860	0.0600
	-0.15	<u>0.5060</u>	0.4050	0.3010	0.2400	0.1830
	-0.10	<u>0.7790</u>	<u>0.6810</u>	<u>0.5630</u>	0.4710	0.3960
	-0.05	<u>0.9320</u>	<u>0.8850</u>	<u>0.8010</u>	<u>0.7410</u>	<u>0.6360</u>
	0	<u>0.9680</u>	<u>0.9450</u>	<u>0.9140</u>	<u>0.8690</u>	<u>0.8080</u>
	0.05	<u>0.9300</u>	<u>0.8910</u>	<u>0.8100</u>	<u>0.7520</u>	<u>0.6700</u>
	0.10	<u>0.7840</u>	<u>0.6940</u>	<u>0.5750</u>	0.4890	0.3960
	0.15	<u>0.5140</u>	0.4150	0.2940	0.2160	0.1450
	0.20	0.2360	0.1480	0.0860	0.0630	0.0360
300	-0.15	0.3960	0.3060	0.2010	0.1490	0.1000
	-0.10	<u>0.7250</u>	<u>0.6430</u>	<u>0.5060</u>	0.4080	0.3230
	-0.05	<u>0.9350</u>	<u>0.8870</u>	<u>0.7910</u>	<u>0.7170</u>	<u>0.6310</u>
	0	<u>0.9760</u>	<u>0.9560</u>	<u>0.9050</u>	<u>0.8470</u>	<u>0.7730</u>
	0.05	<u>0.9280</u>	<u>0.8930</u>	<u>0.7930</u>	<u>0.7120</u>	<u>0.6320</u>
	0.10	<u>0.7250</u>	<u>0.6350</u>	<u>0.5150</u>	0.4020	0.3180
	0.15	0.3970	0.3100	0.1930	0.1290	0.0960

- Dans le cas asymptotique, nous retenons la loi a priori uniforme comme étant la plus appropriée

lorsque le processus généré est du modèle M_1 .

- Dans le cas asymptotique, les lois a priori uniforme, triangulaire et T2Mexp1 favorisent beaucoup plus le modèle M_1 lorsque le processus généré est le modèle M_2 (Exemple: les proportions sont respectivement: 0.7250, 0.6350, 0.5150 pour $n=300$ et $\rho = 0.1$).

- Pour les grandes tailles d'échantillons ($n \geq 200$), lorsque le processus généré est le modèle M_2 , les lois a priori T2Mexp2 et N_{Tr} sont les plus performantes pour le test que les autres lois considérées à partir de $\rho = 0.1$.

Les figures suivantes représentent les allures des fonctions de répartitions empiriques de $P(M_1/X)$ des lois a priori considérées.

Cas des petits échantillons ($n=10$)

Fonctions de repartitions empiriques de $P(M_1=X)$.

Lorsque le processus généré est du modèle M_1 , la fonction de répartition empirique de $P(M_1=X)$ de la loi a priori NTr montre une croissance très élevée autour de la médiane et elle est la plus rapide que les autres fonctions a atteindre la valeur 1.

Cas des grands échantillons (n=100).

Fonctions de répartitions empiriques de $P(M_1 = X)$ pour n=100.

D'après les allures des fonctions de répartitions empiriques associées aux lois a priori NTr et T2Mexp2, nous constatons une croissance élevée autour de la médiane. Ces fonctions atteignent la valeur 1 plus rapidement que les autres fonctions associées aux autres lois a priori lorsque le processus généré est du modèle M_2 .

Lorsque le processus généré est du modèle M_1 , la fonction de répartition empirique de la loi a priori uniforme montre une croissance élevée autour de la médiane, elle est plus rapide que les autres lois a priori à atteindre la valeur 1.

7.2 La puissance des tests. Nous

considérons les statistiques de tests d'indépendance suivantes .

- La statistique des moindres carrés notée $\hat{\rho}_{LSE}$ soit

$$\hat{\rho}_{LSE} = \frac{\sum_{t=2}^n X_t X_{t-1}}{\sum_{t=2}^n X_{t-1}^2}$$

- La statistique d'Hurwicz (1950) notée $\hat{\rho}_{HUR}$ qui s'écrit

$$\hat{\rho}_{HUR} = \text{Med}\left\{\frac{X_2}{X_1}, \frac{X_3}{X_2}, \dots, \frac{X_{n-1}}{X_n}\right\}$$

Cet estimateur représente la médiane des rapports des observations consécutives.

- La statistique de Haddad (2000) notée $\hat{\rho}_{MED}$ donnée par

$$\hat{\rho}_{MED} = \frac{\text{Med}\{X_1 X_2, X_2 X_3, \dots, X_{n-1} X_n\}}{\text{Med}\{X_1^2, X_2^2, \dots, X_{n-1}^2\}}$$

qui représente la médiane normalisée des produits d'observations consécutives.

Dans cette partie, nous comparons nos tests bayésiens (basés sur les lois a priori uniforme, triangulaire, T2Mexp1, T2Mexp2 et N_{Tr}) aux tests classiques au moyen de la puissance dans le cas des grands échantillons uniquement.

Les valeurs simulées du seuil et puissance des différents tests considérés (tests bayésiens et classiques) sont données dans le tableau suivant:

	ρ	Lois a priori pour ρ					test LSE	test HUR	test MED
		Uniforme	Triangulaire	T2Mexp1	T2Mexp2	N_{Tr}			
α	0	0.039	0.068	0.122	0.164	0.201	0.05	0.05	0.05
$1 - \beta$	0.05	0.067	0.118	0.182	0.222	0.242	0.076	0.060	0.062
$1 - \beta$	0.10	0.161	0.217	0.300	0.358	0.404	0.165	0.093	0.093
$1 - \beta$	0.15	0.281	0.388	0.492	0.558	0.591	0.318	0.156	0.151
$1 - \beta$	0.20	0.493	0.597	0.681	0.728	0.763	0.509	0.247	0.235
$1 - \beta$	0.25	0.682	0.768	0.844	0.872	0.886	0.708	0.356	0.341
$1 - \beta$	0.30	0.844	0.889	0.933	0.952	0.961	0.848	0.483	0.463

Table 4. Les valeurs de la puissance des différents tests, n=100.

- Le seuil du test bayésien n'est pas une valeur fixée comme le montre la table 4. Il dépend de la loi a priori choisie.
- Les tests bayésiens considérés excepté celui de la loi a priori uniforme sont plus puissants que les tests classiques (tests de LSE, HUR et MED).
- Nous constatons que les plus grandes puissances sont enregistrées pour les tests bayésiens basés sur les lois a priori N_{Tr} et T2Mexp2.
- Le test bayésien basé sur la loi a priori uniforme est plus puissant que les tests classiques d'Hurwicz (HUR) et Haddad (MED) et moins puissant que le test basé sur l'estimateur des moindres carrés (LSE).

Chapitre 4:

Test bayésien de unit root sous contamination.

Introduction.

L'analyse des séries chronologiques peut être affectée sensiblement par une contamination dans la structure du modèle. Le changement dans la structure du modèle peut être provoqué par la présence d'une rupture (break) dans la tendance de la série ou bien la présence des données aberrantes (outliers). De plus en plus, on rencontre des travaux liés à l'étude de l'effet des événements économiques majeurs, qualifiés d'outliers, tels que les chocs pétroliers, les guerres, les crises financières, les changements de régime politique, etc, sur les modèles de séries temporelles. Il a été enregistré que le changement dans la structure du modèle biaise de manière significative les tests classiques de la racine unitaire (voir Perron 1992, 1994, Perron et Vogelsang, 1998, Lumsdaine et al, 1997 et Ben David et al, 1997).

Ce chapitre traite le test de la racine unitaire par l'approche bayésienne en présence de deux types de contamination: le premier type suppose l'existence d'un changement structurel dans la moyenne du processus (Marriott et Newbold, 2000), le deuxième type concerne la présence des outliers additifs (AO) dans le modèle AR(1).

L'intérêt d'une telle étude est d'évaluer l'impact de ces deux types de contamination sur les conclusions du test bayésien de la racine unitaire.

Partie I.

Test bayésien de unit root en présence d'une structure de rupture.

L'analyse des changements structurels dans les séries chronologiques a suscité un grand intérêt au cours des dernières années après l'apparition de l'article de Perron (1989). Perron (1989) montre que l'évidence de l'existence d'une racine unitaire dans plusieurs variables macro-économiques pourrait être due à la présence d'un important changement structurel dans la tendance des séries qui est ignoré. On enregistre deux volets liés à l'étude des changements structurels dans les séries temporelles. Le premier concerne la détection des points de rupture dans la série et l'estimation de leur nombre. Le second est lié à l'étude de la présence d'une racine unitaire tenant compte des ruptures structurelles possibles.

Dans la littérature actuelle, on trouve un bon nombre de travaux qui traitent les modèles de séries temporelles avec break ou plusieurs breaks dans la tendance, la moyenne ou la variance (voir Cristiano, 1992, Zivot et Andrews, 1992, Lumsdaine et Papell, 1997, Ben David et al, 1997, Wang et Zivot, 2000). Ce deuxième volet fait l'objet de notre intérêt en l'abordant par l'approche bayésienne.

Dans cette partie, nous considérons le modèle AR(1) sans tendance avec un break dans la constante (break au sens de la moyenne: Marriott et Newbold, 2000) qu'on suppose ignoré. Ici, nous élargissons l'étude établie par Marriott et Newbold (2000) qui considèrent comme loi a priori la loi uniforme à d'autres lois a priori. Cette étude nous permettra de voir l'effet d'une telle rupture sur les conclusions du test et d'observer parmi les lois a priori considérées laquelle est la plus appropriée pour le test.

1 Structures de break dans un AR(1).

Plusieurs auteurs abordent les différents types de structures de break dans les séries chronologiques (voir Perron, 1994, Perron et Vogelsong, 1998, Wang et Zivot, 2000, ...).

Nous définissons dans ce paragraphe les différentes structures de ruptures liés à un seul break dans le modèle AR(1).

1.1 Changement structurel dans la tendance.

Soit le processus généré par l'équation:

$$Y_t = \mu_t + b_t t + \rho Y_{t-1} + \sigma \epsilon_t \quad , t = 1, \dots, n \quad (1-1)$$

où les ϵ_t sont des variables aléatoires *i.i.d* de loi $N(0, 1)$, $|\rho| < 1$.

A titre d'exemple, si on considère le cas d'une seule rupture, le break concerne les deux paramètres μ_t et b_t à l'instant $t = k$ (connu ou inconnu) avec $k \leq n$.

$$\mu_t = \begin{cases} \mu_1 & \text{si } 1 \leq t \leq k \\ \mu_2 & \text{si } k < t \leq n \end{cases}$$

et

$$b_t = \begin{cases} b_1 & \text{si } 1 \leq t \leq k \\ b_2 & \text{si } k < t \leq n \end{cases}$$

Notons que dans cette structure de changement dans la tendance, la variance du processus ne change pas (voir Wang et Zivot, 2000).

1.2 Changement structurel dans la moyenne et la variance. On considère le processus suivant:

$$Y_t = \mu_t + \rho Y_{t-1} + s_t \epsilon_t \quad , t = 1, \dots, n \quad (1-2)$$

avec ϵ_t sont des variables *i.i.d* de loi $N(0, 1)$ et $|\rho| < 1$, où il y a absence donc de tendance.

Le break sera donc effectué au niveau des deux paramètres μ_t et s_t à l'instant $t = k$ avec $k \leq n$ (Wang et Zivot, 2000).

$$\mu_t = \begin{cases} \mu_1 & \text{si } 1 \leq t \leq k \\ \mu_2 & \text{si } k < t \leq n \end{cases}$$

et

$$s_t = \begin{cases} \sigma_1 & \text{si } 1 \leq t \leq k \\ \sigma_2 & \text{si } k < t \leq n \end{cases}$$

1.3 Changement structurel dans la moyenne.

Soit le modèle autorégressif sans tendance suivant:

$$Y_t = \mu_t + \rho Y_{t-1} + \sigma \epsilon_t \quad , t = 1, \dots, n \quad (1-3)$$

avec ϵ_t sont des variables *i.i.d* de loi $N(0, 1)$ et $|\rho| < 1$ et $\mu_t = \mu, \forall t$.

Le processus (Y_t) est stationnaire de moyenne $E(Y_t) = \mu_t$ et de variance $V(Y_t) = \sigma^2$.

Si on considère une rupture à l'instant $t = k$ qui touche seulement le paramètre μ_t (Marriott et Newbold, 2000) en posant:

$$\mu_t = \begin{cases} \mu + \mu_1 & \text{si } 1 \leq t \leq k \\ \mu + \mu_2 & \text{si } k < t \leq n \end{cases}$$

La moyenne du processus s'écrit

$$E(Y_t) = \begin{cases} \mu + \mu_1 & \text{si } t \leq k \\ \mu + \mu_2 & \text{si } t > k \end{cases}$$

2 Le test de la racine unitaire en présence d'un break.

Soit le modèle autorégressif stationnaire d'ordre 1:

$$X_t - \mu = \rho(X_{t-1} - \mu) + \epsilon_t, \quad (1 - 4)$$

où les ϵ_t sont des variables aléatoires *i.i.d* de loi $N(0, 1)$ et $|\rho| < 1$.

Lorsqu'il y'a absence de rupture dans le modèle, on peut exprimer le modèle (1 - 4) en fonction des différences

$$W_t = X_t - X_{t-1}, \quad t = 1, \dots, n$$

Si on suppose qu'on a une structure de rupture dans le modèle AR(1) de type considérée par Marriott et Newbold (2000) en posant

$$Y_t = \begin{cases} \mu_1 + X_t & \text{si } t \leq k \\ \mu_2 + X_t & \text{si } t \geq k + 1 \end{cases} \quad (1 - 5)$$

où k est l'instant de la rupture.

Celle-ci peut être vu comme une rupture dans la moyenne du processus (Y_t) .

Remarques.

1. Si on pose $W'_t = Y_t - Y_{t-1}$, $\forall t = 1, \dots, n$, et $\delta = \mu_2 - \mu_1$.
alors, on peut écrire

$$W'_t = \delta Z_t + (X_t - X_{t-1}) = \delta Z_t + W_t, \quad t = 1, \dots, n \quad (1-6)$$

avec

$$Z_t = \begin{cases} 1 & \text{si } t = k + 1 \\ 0 & \text{sinon} \end{cases}$$

2. Comme l'indique l'expression (1-6), la présence d'une telle rupture dans modèle (1-4) se traduit dans ce cas par la présence d'un outlier additif dans le processus W'_t de magnitude $\delta = \mu_2 - \mu_1$ à l'instant $t = k + 1$.

2.1 Présentation des modèles.

Soit les quatre modèles considérés par Marriott et Newbold (2000):

$$M_1 : \quad W'_t = \epsilon_t$$

$$M_2 : \quad W'_t - \rho W'_{t-1} = \epsilon_t - \epsilon_{t-1}$$

$$M_3 : \quad W'_t - \delta Z_t = \epsilon_t$$

$$M_4 : \quad (W'_t - \delta Z_t) - \rho(W'_{t-1} - \delta Z_{t-1}) = \epsilon_t - \epsilon_{t-1}$$

Les deux modèles M_1 et M_3 étant des modèles de marche aléatoire. Quand aux modèles M_2 et M_4 représentent les modèles stationnaires. Les modèles M_3 et M_4 sont les modèles de rupture en présence d'une structure de rupture de type (1-5).

Notre intérêt porte sur la comparaison du modèle M_1 et M_2 uniquement lorsque une telle rupture se présente dans le modèle AR(1) alors qu'elle est ignorée.

2.2 Fonctions de vraisemblance des modèles.

On se donne un échantillon $W'_t = (W'_1, \dots, W'_n)$, les expressions des fonctions de vraisemblance des modèles M_i , $i=1, \dots, 4$ sont données par Newbold (1974).

La fonction de vraisemblance du modèle M_1 en fonction de W'_t s'écrit:

$$P(W'/\sigma, M_1) = \left(\frac{1}{2\pi\sigma^2}\right)^{n/2} \exp\left\{-\frac{1}{2\sigma^2} \sum_{i=1}^n W_t'^2\right\}$$

pour le modèle M_2

$$P(W'/\rho, \sigma, M_2) = \left(\frac{1}{2\pi\sigma^2}\right)^{n/2} A^{\frac{1}{2}} \exp\left\{-\frac{1}{2\sigma^2} \sum_{i=1}^{n+2} u_t^2\right\}.$$

où

$$A = [1 + n(1 - \rho)(1 + \rho)^{-1}]^{-1},$$

$$u_1 = -AC,$$

$$u_2 = \rho(1 - \rho^2)^{-\frac{1}{2}} AC,$$

$$u_3 = W'_1 - (1 + \rho)^{-1} AC,$$

$$u_t = W'_{t-2} + (1 - \rho) \sum_{i=1}^{t-3} W'_j (1 + \rho)^{-1} AC; \quad t = 4, \dots, n + 2.$$

avec

$$C = (1 - \rho) \sum_{t=1}^n W'_t + (1 - \rho)^2 \sum_{i=1}^n (n - t) W'_t.$$

La fonction de vraisemblance du modèle M_3 s'écrit

$$P(W'/\delta, \sigma, k, M_3) = \left(\frac{1}{2\pi\sigma^2}\right)^{n/2} \exp\left\{-\frac{1}{2\sigma^2} \left[\sum_{t \neq k+1} W_t'^2 + (W'_{k+1} - \delta)^2 \right]\right\}$$

Pour le modèle M_4 ,

$$P(W'/\rho, \delta, \sigma, k, M_4) = \left(\frac{1}{2\pi\sigma^2}\right)^{n/2} A^{1/2} \exp\left\{-\frac{1}{2\sigma^2} \sum_{t=1}^{n+2} (u_t^2 + \delta d_t)^2\right\}$$

où

$$d_1 = AC_2,$$

$$d_2 = -\rho(1 - \rho^2)^{-\frac{1}{2}}d_1,$$

$$d_3 = (1 + \rho)^{-1}AC_2 - X_1,$$

$$d_t = -(X_{t-2} - \rho X_{t-3}) + d_{t-1}; \quad t = 4, \dots, n + 2.$$

avec

$$C_2 = (1 - \rho)[1 + (1 - \rho)(n - k + 1)].$$

2.3 Les densités jointes intégrées des modèles.

Les densités a priori jointes des différents modèles s'écrivent:

$$M_1 : P(\sigma/M_1)$$

$$M_2 : P(\rho, \sigma/M_2) = P(\rho)p(\sigma)$$

$$M_3 : P(\delta, \sigma, k/M_3) = p(\delta/\sigma)p(\sigma)p(k)$$

$$M_4 : P(\rho, \delta, \sigma, k/M_4) = p(\rho)p(\delta/\sigma)p(\sigma)p(k)$$

Marriott et Newbold (2000) considèrent les lois a priori suivantes:

La loi a priori non informative pour σ , $p(\sigma) = 1/\sigma$

la densité a priori discrète uniforme pour k , $p(k) = 1/n$ pour $k = 1, \dots, n$.

La loi a priori de δ conditionnelle à σ , $p(\delta/\sigma) \sim N(0, L^2\sigma^2)$.

La densité jointe intégrée des paramètres dans le modèle M_1 s'écrit:

$$P(W'/M_1) = (2\pi)^{-\frac{n}{2}} \Gamma\left(\frac{n}{2}\right) 2^{\frac{n-2}{2}} \left[\sum_{t=1}^n W_t'^2\right]^{-\frac{n}{2}}$$

Dans le modèle M_2 est donnée comme suit:

$$P(W'/M_2) = (2\pi)^{-\frac{n}{2}} \Gamma\left(\frac{n}{2}\right) 2^{\frac{n-2}{2}} \int_{-1}^1 A^{\frac{1}{2}} \left[\sum_{t=1}^{n+2} u_t^2\right]^{-\frac{n}{2}} p(\rho) d\rho$$

La densité jointe intégrée de (δ, σ, k, W') dans le modèle M_3 s'obtient comme suit:

$$P(W'/M_3) = \frac{\pi^{-n/2}}{2nLh^{1/2}} \Gamma\left(\frac{n}{2}\right) \sum_{k=1}^{n-1} \left[\sum_{t=1}^n W_t'^2 - h^{-1}W_{k+1}'^2\right]^{-\frac{n}{2}}$$

avec $h = 1 - L^2$

La densité jointe intégrée de (δ, σ, k, W') pour le modèle M_4 s'obtient comme suit:

$$P(W'/M_4) = \frac{\pi^{-n/2}}{2nLh^{1/2}} \Gamma\left(\frac{n}{2}\right) \sum_{k=0}^{n-1} \int_{-1}^1 A^{1/2} \left[\sum_{t=1}^{n+2} u_t^2 - h^{-1} \left(\sum_{t=1}^{n+2} u_t d_t \right)^2 \right]^{-\frac{n}{2}} p(\rho) d\rho.$$

3 Etude de Monte-Carlo.

Nous nous intéressons dans cette partie à l'impact engendré par la présence de cette rupture qu'on suppose ignorée sur les conclusions du test bayésien de la racine unitaire basé sur les différentes lois a priori étudiées en chapitre 2.

Nous considérons le modèle M_1 et M_2 en présence de la structure de rupture définie par l'expression (1-5) avec

$$P(M_1) = P(M_2) = \frac{1}{2}.$$

Soit le modèle AR(1) stationnaire suivant:

$$u_t = \rho u_{t-1} + \epsilon_t, \quad (1-7)$$

où les ϵ_t sont des variables aléatoires *i.i.d* de loi $N(0, 1)$, $|\rho| < 1$.

En présence de la structure de rupture suivante:

$$Y_t = \begin{cases} u_t & \text{si } t \leq k \\ \delta + u_t & \text{si } t \geq k + 1 \end{cases}$$

ici, $\delta = \mu_2$ sans perte de généralité.

On pose $W'_t = Y_t - Y_{t-1}$, le modèle s'écrit alors:

$$\begin{cases} W'_t = \delta Z_t + (u_t - u_{t-1}) \\ u_t = \rho u_{t-1} + \epsilon_t \end{cases} \quad \text{avec} \quad Z_t = \begin{cases} 1 & \text{si } t = k + 1 \\ 0 & \text{sinon} \end{cases}$$

A base de 1000 répétitions, nous simulons la proportion des échantillons pour lesquels la probabilité $P(M_1/W') > 0.5$ en fonction de δ et k qu'on note $P_{\delta,k}$ pour les tailles d'échantillons $n=10$ et $n=100$. Nous considérons comme loi a priori pour ρ : la loi uniforme et triangulaire sur l'intervalle $[-1, 1]$, la loi Beta1, la loi Beta2, la loi T1Mexp1 et la loi T1Mexp2.

Remarque:

Les simulations effectuées dans cette partie ont révélé que la loi a priori triangulaire (respectivement Beta1, Beta2) se comporte de manière similaire avec la loi a priori uniforme (respectivement T1Mexp1, T1Mexp2). Ainsi, nous mettons dans l'annexe1 les tables concernant les lois a priori triangulaire, Beta1 et Beta2. De plus, nous restreignons l'interprétation des résultats aux lois a priori uniforme, T1Mexp1 et T1Mexp2.

3.1 Cas des petits échantillons .**3.1.1 Loi a priori uniforme .**

Table I.1. Proportion d'échantillons pour lesquels $P(M_1/W') > 0.5$.
La loi a priori uniforme, $n=10$.

ρ	δ										
	0.0	1.0	1.5	2.0	2.5	3.0	3.5	4.0	5.0	10	20
	$k = 5$										
0.0	0.09	0.25	0.46	<u>0.66</u>	<u>0.81</u>	<u>0.90</u>	<u>0.96</u>	<u>0.98</u>	<u>1.00</u>	<u>1.00</u>	<u>1.00</u>
0.1	0.14	0.31	<u>0.51</u>	<u>0.70</u>	<u>0.83</u>	<u>0.92</u>	<u>0.96</u>	<u>0.99</u>	<u>1.00</u>	<u>1.00</u>	<u>1.00</u>
0.3	0.26	0.43	<u>0.62</u>	<u>0.75</u>	<u>0.87</u>	<u>0.94</u>	<u>0.98</u>	<u>0.99</u>	<u>1.00</u>	<u>1.00</u>	<u>1.00</u>
0.5	0.44	<u>0.57</u>	<u>0.70</u>	<u>0.79</u>	<u>0.87</u>	<u>0.93</u>	<u>0.98</u>	<u>0.99</u>	<u>1.00</u>	<u>1.00</u>	<u>1.00</u>
0.7	0.63	<u>0.70</u>	<u>0.75</u>	<u>0.82</u>	<u>0.87</u>	<u>0.92</u>	<u>0.95</u>	<u>0.98</u>	<u>1.00</u>	<u>1.00</u>	<u>1.00</u>
0.9	0.79	<u>0.81</u>	<u>0.83</u>	<u>0.84</u>	<u>0.87</u>	<u>0.90</u>	<u>0.92</u>	<u>0.94</u>	<u>0.98</u>	<u>1.00</u>	<u>1.00</u>
1.0	0.87	<u>0.87</u>	<u>0.86</u>	<u>0.86</u>	<u>0.87</u>	<u>0.89</u>	<u>0.90</u>	<u>0.92</u>	<u>0.94</u>	<u>1.00</u>	<u>1.00</u>
	$k = 2$										
0.0	0.08	0.22	0.36	<u>0.55</u>	<u>0.72</u>	<u>0.83</u>	<u>0.91</u>	<u>0.96</u>	<u>0.99</u>	<u>1.00</u>	<u>1.00</u>
0.1	0.12	0.27	0.42	<u>0.61</u>	<u>0.77</u>	<u>0.87</u>	<u>0.93</u>	<u>0.97</u>	<u>1.00</u>	<u>1.00</u>	<u>1.00</u>
0.3	0.25	0.41	<u>0.57</u>	<u>0.74</u>	<u>0.84</u>	<u>0.92</u>	<u>0.96</u>	<u>0.99</u>	<u>1.00</u>	<u>1.00</u>	<u>1.00</u>
0.5	0.44	<u>0.58</u>	<u>0.70</u>	<u>0.79</u>	<u>0.89</u>	<u>0.95</u>	<u>0.98</u>	<u>0.99</u>	<u>1.00</u>	<u>1.00</u>	<u>1.00</u>
0.7	0.63	<u>0.71</u>	<u>0.76</u>	<u>0.83</u>	<u>0.89</u>	<u>0.94</u>	<u>0.97</u>	<u>0.98</u>	<u>1.00</u>	<u>1.00</u>	<u>1.00</u>
0.9	0.79	<u>0.81</u>	<u>0.83</u>	<u>0.86</u>	<u>0.87</u>	<u>0.90</u>	<u>0.93</u>	<u>0.94</u>	<u>0.97</u>	<u>1.00</u>	<u>1.00</u>
1.0	0.86	<u>0.86</u>	<u>0.86</u>	<u>0.88</u>	<u>0.88</u>	<u>0.90</u>	<u>0.92</u>	<u>0.94</u>	<u>0.96</u>	<u>1.00</u>	<u>1.00</u>

- Nous remarquons que lorsque $\delta \rightarrow \infty$, la proportion $P_{\delta,k} \rightarrow 1 \forall \rho$. Ce qui veut dire que plus δ augmente, plus l'hypothèse nulle est favorisée.
- Nous constatons que le test est affecté. En effet, la conclusion du test change (Exemple: pour $\rho = 0.5$, $\delta = 1$, $\forall k$).

3.1.2 Loi a priori T1Mexp1.

Table I.2. Proportion d'échantillons pour lesquels $P(M_1/W') > 0.5$.
La loi a priori T1Mexp1, n=10.

ρ	δ										
	0.0	1.0	1.5	2.0	2.5	3.0	3.5	4.0	5.0	10	20
$k = 5$											
0.0	0.00	0.04	0.07	0.13	0.22	0.34	0.44	<u>0.53</u>	<u>0.69</u>	<u>0.98</u>	<u>1.00</u>
0.1	0.01	0.04	0.09	0.16	0.26	0.37	0.47	<u>0.56</u>	<u>0.72</u>	<u>0.98</u>	<u>1.00</u>
0.3	0.03	0.07	0.12	0.22	0.33	0.44	<u>0.54</u>	<u>0.63</u>	<u>0.76</u>	<u>0.99</u>	<u>1.00</u>
0.5	0.05	0.11	0.18	0.30	0.42	<u>0.50</u>	<u>0.59</u>	<u>0.67</u>	<u>0.79</u>	<u>1.00</u>	<u>1.00</u>
0.7	0.15	0.21	0.30	0.39	0.48	<u>0.54</u>	<u>0.62</u>	<u>0.67</u>	<u>0.78</u>	<u>0.99</u>	<u>1.00</u>
0.9	0.39	0.41	0.45	0.48	<u>0.53</u>	<u>0.58</u>	<u>0.62</u>	<u>0.66</u>	<u>0.74</u>	<u>0.95</u>	<u>1.00</u>
1.0	<u>0.57</u>	<u>0.56</u>	<u>0.56</u>	<u>0.56</u>	<u>0.57</u>	<u>0.58</u>	<u>0.61</u>	<u>0.63</u>	<u>0.68</u>	<u>0.88</u>	<u>1.00</u>
$k = 2$											
0.0	0.00	0.02	0.04	0.09	0.17	0.26	0.35	0.44	<u>0.59</u>	<u>0.96</u>	<u>1.00</u>
0.1	0.01	0.03	0.05	0.11	0.21	0.28	0.39	0.48	<u>0.62</u>	<u>0.97</u>	<u>1.00</u>
0.3	0.02	0.04	0.09	0.16	0.27	0.36	0.46	<u>0.54</u>	<u>0.68</u>	<u>0.99</u>	<u>1.00</u>
0.5	0.05	0.10	0.17	0.25	0.34	0.44	<u>0.54</u>	<u>0.61</u>	<u>0.74</u>	<u>0.99</u>	<u>1.00</u>
0.7	0.13	0.22	0.27	0.35	0.44	<u>0.52</u>	<u>0.58</u>	<u>0.64</u>	<u>0.76</u>	<u>0.98</u>	<u>1.00</u>
0.9	0.37	0.40	0.44	0.47	<u>0.52</u>	<u>0.56</u>	<u>0.60</u>	<u>0.63</u>	<u>0.71</u>	<u>0.93</u>	<u>1.00</u>
1.0	<u>0.54</u>	<u>0.53</u>	<u>0.55</u>	<u>0.57</u>	<u>0.58</u>	<u>0.59</u>	<u>0.61</u>	<u>0.63</u>	<u>0.67</u>	<u>0.87</u>	<u>0.98</u>

- Le test basé sur la loi a priori T1Mexp1 est robuste pour de petites variations de δ ($0 < \delta \leq 2$).
- Plus δ augmente, ($\delta \geq 2.5$), la conclusion du test change (Exemple: $\delta \longrightarrow \infty$, la proportion $P_{\delta,k} \longrightarrow 1$).

Remarque:

Nous remarquons que l'instant k du break ne joue pas de rôle sur les conclusions du test dans le cas des petits échantillons.

3.2 Cas des grands échantillons .

3.2.1 Loi a prioï uniforme.

Table I.3. Proportion d'échantillons pour lesquels $P(M_1/W') > 0.5$.
La loi a priori uniforme, n=100.

ρ	δ										
	0.0	1.0	1.5	2.0	2.5	3.0	3.5	4.0	5.0	10	20
$k = 50$											
0.0	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.03	0.53	1.00	1.00
0.1	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.09	0.71	1.00	1.00
0.3	0.00	0.00	0.00	0.00	0.00	0.01	0.07	0.34	0.93	1.00	1.00
0.5	0.00	0.00	0.00	0.00	0.01	0.09	0.35	0.68	0.98	1.00	1.00
0.7	0.00	0.01	0.02	0.09	0.24	0.47	0.69	0.86	0.98	1.00	1.00
0.9	0.62	<u>0.65</u>	<u>0.69</u>	<u>0.74</u>	<u>0.79</u>	<u>0.83</u>	<u>0.87</u>	<u>0.92</u>	<u>0.96</u>	<u>1.00</u>	<u>1.00</u>
1.0	0.97	<u>0.98</u>	<u>0.98</u>	<u>0.98</u>	<u>0.98</u>	<u>0.97</u>	<u>0.97</u>	<u>0.98</u>	<u>0.98</u>	<u>0.98</u>	<u>1.00</u>
$k = 5$											
0.0	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.35	1.00
0.1	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.53	1.00
0.3	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.82	1.00
0.5	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.97	1.00
0.7	0.00	0.00	0.00	0.01	0.01	0.01	0.03	0.06	0.16	1.00	1.00
0.9	0.62	<u>0.62</u>	<u>0.65</u>	<u>0.66</u>	<u>0.69</u>	<u>0.73</u>	<u>0.76</u>	<u>0.80</u>	<u>0.88</u>	<u>1.00</u>	<u>1.00</u>
1.0	0.98	<u>0.97</u>	<u>0.97</u>	<u>0.97</u>	<u>0.97</u>	<u>0.97</u>	<u>0.98</u>	<u>0.98</u>	<u>0.98</u>	<u>0.97</u>	<u>1.00</u>

- Pour les petites valeurs de k ($k = 5$), la décision du test ne change que lorsque δ est assez grand ($\delta > 5$).
- Dans le cas contraire ($k = 50$), le test est plus sensible au break ($\delta \geq 3.5$).
- Nous remarquons que lorsque $\delta \longrightarrow \infty$, $P_{\delta,k} \longrightarrow 1 \quad \forall \rho$.

3.2.2 Loi a priori T1Mexp1.

Table I.4. Proportion d'échantillons pour lesquels $P(M_1/W') > 0.5$.
La loi a priori T1Mexp1, n=100.

ρ	δ										
	0.0	1.0	1.5	2.0	2.5	3.0	3.5	4.0	5.0	10	20
$k = 50$											
0.0	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.01	<u>1.00</u>	<u>1.00</u>
0.1	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.02	<u>1.00</u>	<u>1.00</u>
0.3	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.13	<u>1.00</u>	<u>1.00</u>
0.5	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.04	0.39	<u>1.00</u>	<u>1.00</u>
0.7	0.00	0.00	0.00	0.00	0.01	0.03	0.10	0.25	<u>0.68</u>	<u>1.00</u>	<u>1.00</u>
0.9	0.19	0.24	0.28	0.35	0.40	<u>0.50</u>	<u>0.58</u>	<u>0.66</u>	<u>0.79</u>	<u>1.00</u>	<u>1.00</u>
1.0	<u>0.90</u>	<u>0.89</u>	<u>0.89</u>	<u>0.90</u>	<u>0.90</u>	<u>0.91</u>	<u>0.91</u>	<u>0.91</u>	<u>0.92</u>	<u>0.94</u>	<u>0.99</u>
$k = 5$											
0.0	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.02	<u>1.00</u>
0.1	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.04	<u>1.00</u>
0.3	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.14	<u>1.00</u>
0.5	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.35	<u>1.00</u>
0.7	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	<u>0.70</u>	<u>1.00</u>
0.9	0.21	0.22	0.23	0.24	0.27	0.30	0.33	0.37	0.45	<u>0.93</u>	<u>1.00</u>
1.0	<u>0.92</u>	<u>0.91</u>	<u>0.90</u>	<u>0.90</u>	<u>0.90</u>	<u>0.90</u>	<u>0.90</u>	<u>0.90</u>	<u>0.90</u>	<u>0.92</u>	<u>0.96</u>

- Nous enregistrons un comportement similaire de l'impact de la valeur de δ et k sur le test basé sur la loi a priori T1Mexp1 que le test basé sur la loi a priori uniforme.

3.2.2 Loi a priori T1Mexp2.

Table I.5. Proportion d'échantillons pour lesquels $P(M_1/W') > 0.5$.
La loi a priori T1Mexp2, n=100.

ρ	δ										
	0.0	1.0	1.5	2.0	2.5	3.0	3.5	4.0	5.0	10	20
$k = 50$											
0.0	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.06	<u>1.00</u>
0.1	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.11	<u>1.00</u>
0.3	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.25	<u>1.00</u>
0.5	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.46	<u>1.00</u>
0.7	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.66	<u>1.00</u>
0.9	0.00	0.00	0.00	0.01	0.01	0.02	0.04	0.06	0.12	0.73	<u>1.00</u>
1.0	<u>0.59</u>	<u>0.59</u>	<u>0.59</u>	<u>0.60</u>	<u>0.59</u>	<u>0.59</u>	<u>0.59</u>	<u>0.60</u>	<u>0.60</u>	<u>0.66</u>	<u>0.82</u>
$k = 5$											
0.0	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	<u>0.67</u>
0.1	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	<u>0.76</u>
0.3	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	<u>0.91</u>
0.5	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.01	<u>0.96</u>
0.7	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.03	<u>0.98</u>
0.9	0.00	0.00	0.00	0.00	0.00	0.01	0.01	0.01	0.02	0.31	<u>0.98</u>
1.0	<u>0.62</u>	<u>0.63</u>	<u>0.63</u>	<u>0.62</u>	<u>0.61</u>	<u>0.61</u>	<u>0.62</u>	<u>0.62</u>	<u>0.63</u>	<u>0.63</u>	<u>0.73</u>

- Pour les valeurs modérées de δ et $\forall k$, la conclusion du test n'est pas affectée (i.e: robuste).
- Pour les grandes valeurs de δ , la conclusion du test est affectée (lorsque $\delta \longrightarrow \infty$, alors $P_{\delta,k} \longrightarrow 1$) ceci $\forall \rho, \forall k$.

Partie II.

Test bayésien de la racine unitaire en présence d'outliers additifs.

Il est assez fréquent que la présence des données aberrantes ou "outliers" affecte sensiblement les procédures classiques d'inférence en séries chronologiques. La présence de ce type de données provoque un changement dans le régime de la série. Le premier type d'outliers dit *outlier additif (AO)* présente une erreur grossière ou erreur de mesure. Il modélise une erreur soudaine et ponctuelle car celle-ci affecte une seule observation. Le deuxième type qui est dit *outlier d'innovation (IO)* présente une rupture dans la structure du modèle qui se produit lentement et s'étend sur un intervalle de temps. Celui-ci n'affecte pas seulement une seule observation mais plutôt toutes les observations ultérieures.

Dans cette partie, nous étudions l'effet que peut produire la présence d'outliers additifs (AO) sur les conclusions du test bayésien de la racine unitaire. L'étude établie dans cette partie concerne le cas asymptotique seulement.

1 Modèles d'outliers en séries chronologiques.

Le problème d'outliers (valeurs aberrantes) en séries chronologiques a été abordé par la première fois par Fox(1972) dans le modèle $AR(p)$. L'auteur classe les outliers en deux types:

Définition 8.1 *Le type I (observation) est telle que une erreur grossière ou erreur de mesure ou de saisie affecte une seule observation.*

Modèle AO:

On peut présenter un outlier additif AO pour un processus autorégressif d'ordre p ,

$$x_t = \sum_{k=1}^p \rho_k x_{t-k} + \epsilon_t \quad t = p+1, \dots, n \quad (2-1)$$

où les ρ_k sont les paramètres autorégressifs, les $\{\epsilon_t\}$ sont des variables indépendantes de loi $N(0, \sigma_\epsilon^2)$. Les observations $\{y_t\}$ sont telles que

$$y_t = \begin{cases} x_t & \text{si } t \neq q \\ x_t + \Delta & \text{si } t = q \end{cases}$$

La valeur aberrante agit sur la série $\{y_t\}$ et non sur la série $\{x_t\}$.

Définition 8.2 *Le type II (innovation) est lorsque un outlier correspond à la situation où une seule "innovation" prend une valeur extrême. Celle-ci n'affecte pas seulement une seule observation mais plutôt toutes les observations ultérieures.*

Modèle IO:

Toujours dans le modèle $AR(p)$ défini par l'expression (2-1), les observations $\{y_t\}$ sont telles que

$$y_t = \sum_{k=1}^p \rho_k x_{t-k} + \epsilon_t + \Delta_t \quad t = p+1, \dots, n$$

où

$$\Delta_t = \begin{cases} 0 & \text{si } t \neq q \\ \Delta & \text{si } t = q \end{cases}$$

Les ρ_k ($k=1, \dots, p$) et $\{\epsilon_t\}$ sont définis précédemment. Ici, les outliers affectent l'observation y_q et aussi la suite des observations $y_{q+1}, y_{q+2}, \dots, y_n$.

Fox (1972) a présenté trois possibilités d'apparition des outliers dans un processus: les outliers présents dans la série sont de type I, les outliers sont de type II et les outliers sont mixtes (mélange des deux types).

2 Test de unit root sous contamination AO.

Nous considérons le modèle autorégressif d'ordre 1 suivant

$$X_t - \mu = \rho(X_{t-1} - \mu) + \epsilon_t, \quad (2-2)$$

où les ϵ_t sont des variables aléatoires *i.i.d* $N(0, 1)$, $|\rho| < 1$ et μ est un paramètre inconnu.

On suppose que suite à des erreurs de mesure, on observe Y_1, \dots, Y_n , au lieu de X_1, \dots, X_n tel que:

$$Y_t = \begin{cases} X_t & \text{si } t \neq k \\ X_t + \Delta & \text{si } t = k \end{cases}$$

où k est l'instant de l'apparition de l'outlier de type AO et Δ sa magnitude.

On pose, $W'_t = Y_t - Y_{t-1}$ et $W_t = X_t - X_{t-1}$, $\forall t = 1, \dots, n$
on peut écrire:

$$W'_t = \Delta Z_t + W_t, \quad t = 1, \dots, n \quad \text{avec} \quad Z_t = \begin{cases} 1 & \text{si} \quad t = k \\ -1 & \text{si} \quad t = k + 1 \\ 0 & \text{sinon.} \end{cases} \quad (2-3)$$

Remarque:

Comme l'indique l'expression (2-3), la présence d'un AO dans la série d'observations Y_1, \dots, Y_n se traduit par la présence de deux outliers additifs aux instants $t = k$ et $t = k + 1$ dans le processus W'_t qui sont de magnitude Δ et $(-\Delta)$ respectivement.

Nous abordons par l'approche bayésienne le test de la racine unitaire

$$H_0 : \rho = 1 \quad \text{contre} \quad H_1 : \rho < 1$$

Comme un problème bayésien de sélection des deux modèles:

$$\begin{cases} M_1 : W'_t = \epsilon_t \\ M_2 : W'_t - \rho W'_{t-1} = \epsilon_t - \epsilon_{t-1} \end{cases}$$

Nous reprenons les différentes lois a priori de ρ considérées au chapitre 2 à savoir la loi uniforme et triangulaire sur $[-1, 1]$, la loi Beta1, la loi Beta2, la loi T1Mexp1 et la loi T1Mexp2.

3 Etude de Monte-Carlo.

Afin d'observer l'impact de la présence d'un AO ou plusieurs AO sur la décision du test, nous simulons la proportion des échantillons pour lesquels $P(M_1/W') > 0.5$ en fonction de Δ qu'on note Pr_Δ en considérant les différentes lois a priori citées ci-dessus. Les simulations effectuées sont à base de 1000 échantillons de taille $n=100$.

- Pour le cas de présence d'un AO, on considère deux instants de son apparition $k = 5$ et $k = 50$.
- Pour le cas de multiples AO (trois AO), on considère qu'ils apparaissent avec la même magnitude Δ aux instants $k_1 = 5$, $k_2 = 20$ et $k_3 = 80$.

Remarque:

Nous avons remarqué que loi a priori triangulaire (respectivement Beta1, Beta2) se comporte de manière similaire avec la loi a priori uniforme (respectivement T1Mexp1, T1Mexp2). Ainsi, nous mettons dans l'annexe2 les tables concernant les lois a priori triangulaire, Beta1 et Beta2. Et nous

restreignons l'interprétation des résultats aux lois uniforme, T1Mexp1 et T1Mexp2.

a) **Cas de la loi a priori uniforme.**

Table II.1. Proportion d'échantillons pour lesquels $P(M_1/W') > 0.5$
Loi uniforme avec un AO.

ρ	Δ										
	0.0	1.0	1.5	2.0	2.5	3.0	3.5	4.0	5.0	10	20
$k = 50$											
0.70	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
0.80	0.08	0.09	0.09	0.09	0.08	0.09	0.09	0.09	0.09	0.06	0.02
0.85	0.30	0.30	0.30	0.30	0.30	0.30	0.30	0.30	0.30	0.27	0.21
0.90	0.62	<u>0.63</u>	<u>0.62</u>	<u>0.62</u>	<u>0.62</u>	<u>0.62</u>	<u>0.62</u>	<u>0.61</u>	<u>0.62</u>	<u>0.64</u>	<u>0.72</u>
0.95	0.89	<u>0.87</u>	<u>0.86</u>	<u>0.86</u>	<u>0.86</u>	<u>0.86</u>	<u>0.86</u>	<u>0.86</u>	<u>0.86</u>	<u>0.89</u>	<u>0.98</u>
1.00	0.97	<u>0.98</u>	<u>0.97</u>	<u>0.97</u>	<u>0.97</u>	<u>0.98</u>	<u>0.98</u>	<u>0.97</u>	<u>0.98</u>	<u>0.98</u>	<u>1.00</u>
$k = 5$											
0.70	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
0.80	0.07	0.07	0.07	0.07	0.07	0.07	0.07	0.07	0.07	0.06	0.01
0.85	0.29	0.29	0.29	0.28	0.29	0.28	0.29	0.28	0.27	0.24	0.18
0.90	0.63	<u>0.62</u>	<u>0.62</u>	<u>0.62</u>	<u>0.62</u>	<u>0.62</u>	<u>0.61</u>	<u>0.61</u>	<u>0.62</u>	<u>0.64</u>	<u>0.75</u>
0.95	0.87	<u>0.86</u>	<u>0.86</u>	<u>0.86</u>	<u>0.86</u>	<u>0.86</u>	<u>0.86</u>	<u>0.87</u>	<u>0.86</u>	<u>0.90</u>	<u>0.96</u>
1.00	0.98	<u>0.98</u>	<u>0.98</u>	<u>0.97</u>	<u>0.98</u>	<u>0.97</u>	<u>0.97</u>	<u>0.97</u>	<u>0.97</u>	<u>0.99</u>	<u>1.00</u>

Table II.2. Proportion d'échantillons pour lesquels $P(M_1/W') > 0.5$
Loi uniforme avec trois AO.

ρ	Δ										
	0.0	1.0	1.5	2.0	2.5	3.0	3.5	4.0	5.0	10	20
0.50	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
0.80	0.08	0.05	0.02	0.01	0.01	0.01	0.01	0.01	0.00	0.00	0.00
0.85	0.31	0.28	0.20	0.13	0.09	0.07	0.05	0.04	0.03	0.01	0.01
0.90	0.65	<u>0.73</u>	<u>0.79</u>	<u>0.84</u>	<u>0.87</u>	<u>0.88</u>	<u>0.91</u>	<u>0.93</u>	<u>0.95</u>	<u>0.94</u>	<u>0.97</u>
0.95	0.90	<u>0.94</u>	<u>0.95</u>	<u>0.98</u>	<u>0.98</u>	<u>0.98</u>	<u>0.98</u>	<u>0.99</u>	<u>0.99</u>	<u>1.00</u>	<u>1.00</u>
1.00	0.98	<u>0.99</u>	<u>0.99</u>	<u>0.99</u>	<u>1.00</u>	<u>1.00</u>	<u>1.00</u>	<u>1.00</u>	<u>1.00</u>	<u>1.00</u>	<u>1.00</u>

Le test de unit root basé sur la loi a priori uniforme résiste à la présence d'outliers additifs. En effet, la décision du test reste inchangée $\forall \Delta$ (voir table II.1 et II.2).

On remarque des deux tables que:

- Lorsque $0.9 \leq \rho \leq 1$, la proportion $Pr_{\Delta} \longrightarrow 1$ quand $\Delta \longrightarrow \infty$.
- Lorsque $\rho > 0.9$, la proportion $Pr_{\Delta} \longrightarrow 0$ quand $\Delta \longrightarrow \infty$.

b) Loi a priori T1Mexp1.

Table II.3. Proportion d'échantillons pour lesquels $P(M_1/W') > 0.5$
Loi T1Mexp1 avec un AO.

ρ	Δ										
	0.0	1.0	1.5	2.0	2.5	3.0	3.5	4.0	5.0	10	20
$k = 50$											
0.80	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
0.85	0.04	0.04	0.03	0.03	0.04	0.04	0.04	0.04	0.03	0.02	0.00
0.90	0.20	0.20	0.20	0.20	0.20	0.20	0.20	0.19	0.19	0.17	0.10
0.95	<u>0.57</u>	<u>0.57</u>	<u>0.56</u>	<u>0.56</u>	<u>0.56</u>	<u>0.56</u>	<u>0.56</u>	<u>0.56</u>	<u>0.57</u>	<u>0.59</u>	<u>0.66</u>
1.00	<u>0.90</u>	<u>0.90</u>	<u>0.90</u>	<u>0.90</u>	<u>0.90</u>	<u>0.90</u>	<u>0.90</u>	<u>0.90</u>	<u>0.90</u>	<u>0.92</u>	<u>0.99</u>
$k = 5$											
0.80	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
0.85	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.01	0.00
0.90	0.18	0.18	0.18	0.18	0.18	0.17	0.17	0.17	0.17	0.15	0.09
0.95	<u>0.58</u>	<u>0.59</u>	<u>0.59</u>	<u>0.58</u>	<u>0.57</u>	<u>0.56</u>	<u>0.56</u>	<u>0.56</u>	<u>0.56</u>	<u>0.56</u>	<u>0.64</u>
1.00	<u>0.90</u>	<u>0.90</u>	<u>0.90</u>	<u>0.90</u>	<u>0.90</u>	<u>0.91</u>	<u>0.91</u>	<u>0.91</u>	<u>0.90</u>	<u>0.90</u>	<u>0.96</u>

Table II.4. Proportion d'échantillons pour lesquels $P(M_1/W') > 0.5$
Loi T1Mexp1 avec trois AO.

ρ	Δ										
	0.0	1.0	1.5	2.0	2.5	3.0	3.5	4.0	5.0	10	20
0.80	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
0.85	0.04	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
0.90	0.25	0.16	0.08	0.05	0.04	0.04	0.03	0.02	0.02	0.00	0.00
0.95	0.63	<u>0.72</u>	<u>0.76</u>	<u>0.81</u>	<u>0.85</u>	<u>0.87</u>	<u>0.90</u>	<u>0.91</u>	<u>0.94</u>	<u>0.94</u>	<u>0.96</u>
1.00	0.94	<u>0.97</u>	<u>0.98</u>	<u>0.99</u>	<u>0.99</u>	<u>0.99</u>	<u>0.99</u>	<u>1.00</u>	<u>1.00</u>	<u>1.00</u>	<u>1.00</u>

- La présence d'un seul ou plusieurs outliers additifs n'affecte pas la décision du test de unit root basé sur la loi a priori T1Mexp1. En effet la décision du test ne change pas $\forall \Delta$ (table II.3 et II.4).

Nous remarquons de la table II.3 et table II.4 que:

- Pour les valeurs $0.95 \leq \rho \leq 1$, la proportion $Pr_{\Delta} \longrightarrow 1$ quand $\Delta \longrightarrow \infty$.
- Lorsque $\rho > 0.95$, la proportion $Pr_{\Delta} \longrightarrow 0$ lorsque $\Delta \longrightarrow \infty$.

c)La loi a priori T1Mexp2.

Table II.5. Proportion d'échantillons pour lesquels $P(M_1/W') > 0.5$
Loi a priori T1Mexp2 avec un AO.

ρ	Δ										
	0.0	1.0	1.5	2.0	2.5	3.0	3.5	4.0	5.0	10	20
$k = 50$											
0.80	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
0.85	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
0.90	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
0.95	0.08	0.08	0.08	0.08	0.08	0.08	0.08	0.07	0.07	0.03	0.00
1.00	<u>0.62</u>	<u>0.61</u>	<u>0.61</u>	<u>0.61</u>	<u>0.62</u>	<u>0.62</u>	<u>0.61</u>	<u>0.62</u>	<u>0.61</u>	<u>0.65</u>	<u>0.79</u>
$k = 5$											
0.80	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
0.85	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
0.90	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
0.95	0.07	0.06	0.06	0.06	0.05	0.05	0.05	0.05	0.05	0.03	0.00
1.00	<u>0.61</u>	<u>0.62</u>	<u>0.62</u>	<u>0.62</u>	<u>0.62</u>	<u>0.62</u>	<u>0.62</u>	<u>0.62</u>	<u>0.62</u>	<u>0.63</u>	<u>0.70</u>

Table II.6. Proportion d'échantillons pour lesquels $P(M_1/W') > 0.5$
Loi a priori T1Mexp2 avec trois AO.

ρ	Δ										
	0.0	1.0	1.5	2.0	2.5	3.0	3.5	4.0	5.0	10	20
0.80	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
0.85	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
0.90	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
0.95	0.08	0.02	0.01	0.00	0.00	0.01	0.00	0.00	0.00	0.00	0.00
1.00	<u>0.71</u>	<u>0.82</u>	<u>0.89</u>	<u>0.91</u>	<u>0.94</u>	<u>0.96</u>	<u>0.96</u>	<u>0.98</u>	<u>0.98</u>	<u>0.99</u>	<u>1.00</u>

Le test de unit root basé sur la loi a priori T1Mexp2 résiste lui aussi à la présence des outliers additifs. La présence de ce type de données n'affecte pas la décision du test (voir table II.5 et II.6).

De la table II.5 et II.6, nous pouvons tirer les deux propriétés suivantes:

i) Lorsque $\rho = 1$, la proportion $Pr_{\Delta} \longrightarrow 1$ quand $\Delta \longrightarrow \infty$.

ii) Lorsque $\rho < 1$, la proportion $Pr_{\Delta} \longrightarrow 0$ quand $\Delta \longrightarrow \infty$.

Remarques.

- L'instant k de l'apparition de la contamination n'influe pas sur la décision des différents tests .
- Les résultats observés sur le test basé sur la loi a priori uniforme (respectivement T1Mexp1, respectivement T1Mexp2) sont obtenus pour le test basé sur la loi a priori triangulaire (respectivement Beta1, respectivement Beta2).

Conclusions .

Des deux propriétés *i)* et *ii)* que vérifie le test basé sur la loi a priori Beta2 et T1Mexp2, nous concluons que nous pouvons corriger le test en rendant ces deux lois a priori appropriées pour les deux hypothèses du test, pas uniquement pour l'hypothèse alternative. Ceci peut se faire en pratique, en provoquant la présence d'une petite proportion d'outliers additive avec une importante magnitude dans les données observées.

Nous pouvons également conclure que les différents tests considérés sont robuste en présence d'outliers additifs dans le sens où la décision des tests ne change pas.

Conclusion générale

L'étude réalisée dans ce mémoire porte essentiellement sur l'inférence bayésienne liée au paramètre du modèle autorégressif d'ordre 1. En particulier, nous nous sommes intéressé au test bayésien de la racine unitaire et le test d'indépendance.

Nous avons proposé un test bayésien de la racine unitaire basé sur la loi a priori exponentielle modifiée notée T1Mexp que nous avons construit. La performance de ce test est établie lorsque le processus généré est stationnaire dans le cas des grands et petits échantillons. L'étude comparative en terme de la puissance a révélé que les tests bayésiens considérés dans ce travail sont plus puissants que les tests classiques de Dickey-Fuller (D-F) et le test du maximum de vraisemblance (MLE).

L'étude des test bayésiens en présence d'un break au sens de la moyenne montre que le test basé sur les lois a priori T1Mexp2 et Beta2 sont les moins affectés dans le sens où leurs conclusions ne changent que lorsque la magnitude δ est assez importante comparés aux autres tests.

Nous avons également constaté que la contamination par la présence d'outliers additifs dans l'échantillon n'influe pas sur les conclusions des différents tests bayésiens de unit root. D'autre part, nous proposons une correction du test de unit root basé sur une des lois a priori suivantes: la loi proposée T1Mexp2 et la loi non symétrique Beta2 en rendant ces deux lois a priori appropriées pour les deux hypothèses du test pas uniquement pour l'hypothèse alternative.

Nous avons abordé le test d'indépendance par l'approche bayésienne par la proposition de quelques lois a priori. Dans le cas des petits échantillons, toutes les lois a priori considérées ont tendance à favoriser l'hypothèse H_0 . Nous retenons la loi a priori normale tronquée N_{Tr} comme étant la plus appropriée pour H_0 . Cependant, nous ne détenons pas encore une loi a priori qui répond favorablement à l'hypothèse H_1 . Dans le cas asymptotique, le test basé sur la loi a priori normale tronquée et la loi exponentielle modifiée T2Mexp2 sont les plus appropriée pour l'hypothèse H_1 . Tandis que la loi a priori uniforme est la plus appropriée pour l'hypothèse H_0 .

Comme perspectives, il serait intéressant de

- Envisager d'autres types de contaminations pour le test bayésien de la racine unitaire, comme par exemple une contamination de type IO.

- Considérer l'étude des tests bayésiens proposés dans ce travail en présence de rupture dans la tendance ou la moyenne et la variance simultanément.
- Considérer un test bayésien d'indépendance pour un modèle AR(1) avec un terme constant et en présence d'outliers.

Bibliographie.

Akaike, H. (1978). "A new look at the bayes procedure ", *Biometrika*. 65, 53-59

Andrews, D. K. W. (1998). "Tests for white noise against alternatives with both seasonal and nonseasonal serial correlation ", *Biometrika*, 85, 3, 727-740.

Banerjee, A. Lumsdaine, R. et Stock, J.H. (1992). "Recursive and sequential tests for the unit root and trend break hypothesis: Theory and international evidence", *Journal of Business and Economics Statistics*, 10, 271-287.

Bartlett, M. S. (1957). "A comment on D. V. Lindley's statistical paradox", *Biometrika*, 44, 533-534.

Basawa, I. V, Huggins, R. M. and Staudte, R. G. (1985). "Robust tests for time series with an application to first order autoregressive process ", *Biometrika*, 72, 3, 559-571.

Bayes, T. (1763). "An essay towards solving a problem in the doctrine of chances", *Phil. Trans. Roy. Soc*, 53, 370-418.

Ben-David, D. Lumsdaine, R. L. and Papell, D. H. (1997). "Unit roots, postwar slowdown and long-run growth: evidence from two structural breaks", *Working paper 6397, national bureau of econometrics research, cambridge, MA*.

Berger, J. (1980). "Statistical decision theory and Bayesian analysis", *Springer Verlag, New York*.

Berger, J. (1985). "Statistical decision theory and bayesian analysis", *Springer-verlag, New York, Berlin, Hei delberg*.

- Berger, J. and Bernardo, J. (1992). "On the development of reference priors", *Journal of Econometrics*, 49 (1991)195-238.
- Berger, J. and Delampady, M. (1987). "Testing precise hypothesis", *Statistical Sciences*, 2, 317-352.
- Berger, J. and Sellke, T. (1987). "Testing of a point null hypothesis: the irreconcilability of significance levels and evidence", *Journal of the American Statistical Association*, 82, 112-122.
- Berger, J. and Yang, R.Y. (1994). "Noninformative priors and bayesian testing for the AR(1) model", *Econometrics Theory*, 10, 461-482.
- Berkoun, Y. Fellag, H. and Zieliński, R. (2003). "Robust Testing Serial Correlation in AR(1) Processes in the Presence of a Single Additive Outlier", *Communication in statistic, theory and methods* vol. 32. N, pp 1527-1540.
- Bernardo, J.M. (1979). "Reference posterior distributions for bayesian inference", *J. Roy. Statist. Soc*, B 41, 113-147.
- Bernardo, J. M. (1980). "A bayesian analysis of classical hypothesis testing", *In Bayesian Statistics*, (J.M. Bernardo, M.H. De Groot, D.V. Lindley and A. F. M. Smith editor), 605-618, University Press, Valencia.
- Casella, G. and Berger, R. (1987). "Reconciling bayesian and frequentist evidence in the one-sided testing problem", *Journal of American Statistical Association*, vol 82, No 397, 106-111.
- Chib, S. (1998). "Estimation and comparison of multiple changepoint models", *Journal of Econometrics*, 86, 221-241.
- Conigliani, C. and O'Hagan, A. (2000). "Sensitivity of the fractional Bayes factor to prior distributions", *The Canadian Journal of Statistics*, 28, 343-352.
- Conigliani, C. and Spezzaferri, F. (2007). "A robust bayesian approach for unit root testing", *Econometric theory*, 23, 430 -463.

- Cristiano, L.J. (1992). "Searching for a break in GNP", *Journal of Business and Economics Statistics*, 10, 237-256.
- Davis, B. (1983). "On brownian slow points", *Z. W. arhsheinlichkeits theori*, 64, 359-367.
- Dejong, D. and Whitman, C. (1991). "Reconsidering trends and Random Works in macro economic time series", *Journal of Monetary Economics*, 28, 221-254.
- Denby, I. and Martin. (1979). "Robust Estimation of the First Order Autoregressive Parameters", *Journal of American Statistical Association*, Vol.74, No. 365, 140-146.
- Dickey, D. A. and Fuller, W. A. (1979). "Distribution of estimators for autoregressive time series with a unit root", *Journal of The American Statistical Association*, 74, 427-431.
- Droesbeke, J-J. Fine, J. Saporta, G. (2002). "Méthodes bayésiennes en statistiques", *Edition-Technip*.
- Durbin, J. and Watson, G.S. (1950). "Testing for serial correlation in least square regression ", *I Biometrika*, 31, 409-428.
- Fischer, R. A.(1956). "Statistical methods and scientific inference", *Idinburgh - Oliver and Boyd*.
- Fox, A. J. (1972). "Outliers in time series", *J. Roy. Stat. Soc. Ser, B*, 34, 350-363.
- Fuller, W. A. (1976). "Introduction to statistical time series", *New York: John Wiley*.
- Haddad, J. N. (2000). "On robuste estimation in the first-order autoregressive processes ", *Communication in statistics, Theory and Methods*, 29 (1), 45-54.
- Hartigan, J. A. (1964). "Invariant prior distribution", *Ann. Math. Statist*, 35, 836-845.
- Hultblad, A. M. et Karlsson, S. (2006). "Detection of simultaneous determination of structural breaks and lag lengths", *Working paper, Stocholm Schooll of Econometrics*.

- Hurwicz, L. (1950). "Least- Squares Bias in Time Series", *Statistical inference in dynamic economic models*, T. C. Koopmas, ed, Wiley.
- Inclàn, C.V. (1993). "Detection of multiple change of variance using posterior odds", *Journal of business and Econometrics statistics*, 11, 289-300.
- Jeffrey A. Millis. (2008). "Resolving some paradoxes and problems with bayesian precise hypothesis testing".
- Jeffreys, H. (1939). "Theory of probability", *Oxford: Clarendon Press*.
- Jeffreys, H. (1961). "Theory of probability", 3rd édition, *University Press, Oxford*.
- Kass, R. and Raftery, A.E. (1995). "Bayes factors", *Journal of the American Statistical Association*, vol 90, No 430, 773-795.
- Koop, G. et Potter, A. E. (2004 a). "Prior elicitation in multiple change-point models ", *Working paper, University of Leicester*.
- Kulberger, R. J. and Lockhart, T. A. (1995). "Tests of independance in time series ", *Journal of time series analysis*, 19, 2, 165-185.
- Laplace, P.S. (1774). "Mémoire sur les probabilités des causes par les événements" *Mém, Acad. R. Sci. Paris*, 6, 621-656.
- Lindley, D.V. (1956). "On a mesure of the information provided by experiment", *Ann. Math. Statist*, 27, 986-1005.
- Lindley, D. V. (1957). "A Statistical Paradox", *Biometrika*, 44, 187-192.
- Lubrano, M. (1995). "Testing for unit roots in Bayesian framework", *Journal of Econometrics*, 69 (1995)81-109.
- Lumsdaine, R. L. et Papell, D. H. (1997). "Multiple tend breaks and the unit roots hypothesis",

review of Economics and Statistics, 79, 212-218.

- Marin, J. M. and Robert, C.P. (2007). "A practical approach to computational bayesian statistics",
- Marriott, J. And Newbold, P. (1998). "Bayesian comparison of ARIMA and stationary ARMA models", *International Statistical Review*, 66, 323-336.
- Marriott, J. And Newbold, P. (2000). "The strength of evidence for unit autoregressive root and structural breaks: A Bayesian perspective", *Journal of Econometrics*, 98 (2000)1-25.
- Meligkotsidou, L., Tzavalis, E and Vrontos, I. D. (2005). "A Bayesian Analysis of Unit Roots and Structural Breaks in the Level and the Error Variance of autoregressive Models", *Working Paper, Queen Mary, University of London*.
- Nelson, C. and Plosser, C. (1982). "Trends and random walks in macroeconomics time series", *Journal of Monetary Econometrics*, 10, "139-162.
- Newbold, P. (1974). "The exact likelihood function for a mixed autoregressive- moving average process", *Biometrika*, 61, 423-426.
- Neyman, J. and Pearson, E.S. (1933). "On the problem of the most efficient tests of statistical hypothesis", *Philosophical Trans. Royal. Soc. London Serie A*, 231, 289-337.
- O'Hagan, A. (1995). "Fractional Bayes Factors for model comparison (with discussion)", *Journal of the Royal Statistical Society, Series B*, 56, 99-118.
- O'Hagan, A. (1997). "Properties of Intrinsic and Fractional Bayes Factor", *Test*, 6, 101-118.
- Pantula, S. Gonzalez-Farias, G. Fuller, W. A. (1994). "A Comparison of unit root test criteria", *Journal of Business and Economics Statistics*, vol 12 No 4.
- Perron, P.(1989). "The great crash, the oil price shock, and the unit root hypothesis", *Econometrica*, 57, 1361-1401

- Perron, P. (1992). "Trend, Unit Root, and Structural Change: A Multi-Country With Historical Data", in *Proceedings of the Business and Economics Statistics Section, American Statistical Association*, pp. 144-149.
- Perron, P. (1994). "Trend, unit root and structural change in macroeconomics time series", in *cointegration for the applied econometrics*, ed. B. B. Rao, New York: macMilan, PP. 113-146.
- Perron, P. (1997). "Further evidence of breaking trend functions in macroeconomic variables", *Journal of econometrics*, 80, 355-385.
- Perron, P. et Vogelsang, T. J. (1998). "Additional tests for a unit root allowing for a break in the trend function in an unknown time", *International Econometric Review*, 39, 1073-1100.
- Phillips, P. C. (1991). "To criticize the critics: an objective bayesian analysis of stochastic trends", *Journal of Applied Econometrics*.
- Robert, C.P. (1992). "L'analyse statistique bayésienne", *Econometrica, Paris*.
- Robert, C.P. (2000). "The bayesian choice", *Springer-verlag, New York, Berlin, Heidelberg*.
- Schotman, P. and Van Dijk, H. K. (1991a). "A Bayesian analysis of the unit root in real exchange rates", *Journal of Econometrics*, 49, 195-238.
- Schotman, P. and Van Dijk, H. K. (1991b). "On Bayesian roots to unit roots", *Journal of Applied Econometrics*, vol. 6, 387-401.
- Schotman, P. and Van Dijk, H. K. (1993). "Posterior analysis of possibly integrated time series with an application to real GNP", in *New Directions in Time Series Analysis part II, IMA, volumes in Mathematics and its Applications*. 46, ed. by P. Caines, J. Geweke, and M. Taqq, pp. 341-361.
- Schwarz, G. (1978). "Estimating the dimension of the model", *The Annals of Statistics*, vol. 14, pp. 456-471.

- Sims, C. A. (1988). "Bayesian skepticism on unit roots econometrics", *Journal of Econometric Dynamic and Control*, 12, 463.
- Spiegelhater, D.J. Best, N.G. Carlin, B.P. et Van der Linde, A. (2002). "Bayesian measures of model complexity and fit (with discussion)", *J. Roy Statist. Soc. B*, 64, pp. 583-640.
- Wang, J. et Zivot, E. (2000). "A bayesian time series model of multiple structural change in level, trend and variance", *Journal of Business and Econometrics Statistics*, 18, 374-386.
- Zivot, E. et Andrews, D. W. K. (1992). "Further evidence of the great crash, the oil-price shock, and the unit root hypothesis", *Journal of Business and Economics Statistics*, 10, 251-270.

ANNEXE

Nous donnons dans **l'annexe 1** les valeurs simulées de la proportion d'échantillons pour lesquels $P(M_1/W') > 0.5$. Les résultats sont basés sur 1000 répétitions et concernent les lois a priori Triangulaire, Beta1 et Beta2, dans le cas de présence d'un break au sens de la moyenne.

Dans **l'annexe 2**, nous présentons les valeurs simulées de $P(M_1/W') > 0.5$ dans le cas de présence d'outliers additifs pour les mêmes lois a priori.

Dans le cas de présence d'un seul AO, nous considérons deux instants de son apparition: $k = 5$ et $k = 50$.

Dans le cas de présence de plusieurs AO (trois AO), on suppose qu'ils apparaissent avec la même magnitude Δ aux instants $k_1 = 5$, $k_2 = 20$ et $k = 80$.

Les simulations effectuées sont à base de 1000 échantillons de taille $n = 100$.

ANNEXE 1

Table 1.1. Proportion d'échantillons pour lesquels $P(M_1/W') > 0.5$.
La loi a priori triangulaire, n=10.

ρ	δ										
	0.0	1.0	1.5	2.0	2.5	3.0	3.5	4.0	5.0	10	20
$k = 5$											
0.0	0.04	0.15	0.31	<u>0.50</u>	<u>0.66</u>	<u>0.80</u>	<u>0.88</u>	<u>0.93</u>	<u>0.99</u>	<u>1.00</u>	<u>1.00</u>
0.1	0.06	0.19	0.36	<u>0.54</u>	<u>0.71</u>	<u>0.82</u>	<u>0.91</u>	<u>0.95</u>	<u>0.99</u>	<u>1.00</u>	<u>1.00</u>
0.3	0.14	0.30	0.46	<u>0.63</u>	<u>0.76</u>	<u>0.86</u>	<u>0.93</u>	<u>0.97</u>	<u>0.99</u>	<u>1.00</u>	<u>1.00</u>
0.5	0.27	0.42	<u>0.56</u>	<u>0.69</u>	<u>0.79</u>	<u>0.87</u>	<u>0.93</u>	<u>0.97</u>	<u>0.99</u>	<u>1.00</u>	<u>1.00</u>
0.7	0.48	<u>0.55</u>	<u>0.65</u>	<u>0.73</u>	<u>0.80</u>	<u>0.86</u>	<u>0.91</u>	<u>0.95</u>	<u>0.99</u>	<u>1.00</u>	<u>1.00</u>
0.9	<u>0.69</u>	<u>0.71</u>	<u>0.74</u>	<u>0.77</u>	<u>0.79</u>	<u>0.83</u>	<u>0.86</u>	<u>0.90</u>	<u>0.95</u>	<u>1.00</u>	<u>1.00</u>
1.0	<u>0.80</u>	<u>0.81</u>	<u>0.79</u>	<u>0.79</u>	<u>0.80</u>	<u>0.82</u>	<u>0.84</u>	<u>0.86</u>	<u>0.90</u>	<u>1.00</u>	<u>1.00</u>
$k = 2$											
0.0	0.04	0.12	0.24	0.38	<u>0.56</u>	<u>0.70</u>	<u>0.82</u>	<u>0.88</u>	<u>0.96</u>	<u>1.00</u>	<u>1.00</u>
0.1	0.05	0.16	0.28	0.44	<u>0.61</u>	<u>0.75</u>	<u>0.85</u>	<u>0.91</u>	<u>0.98</u>	<u>1.00</u>	<u>1.00</u>
0.3	0.13	0.27	0.41	<u>0.57</u>	<u>0.72</u>	<u>0.82</u>	<u>0.90</u>	<u>0.95</u>	<u>0.99</u>	<u>1.00</u>	<u>1.00</u>
0.5	0.29	0.41	<u>0.55</u>	<u>0.68</u>	<u>0.78</u>	<u>0.86</u>	<u>0.93</u>	<u>0.96</u>	<u>0.99</u>	<u>1.00</u>	<u>1.00</u>
0.7	0.48	<u>0.57</u>	<u>0.65</u>	<u>0.73</u>	<u>0.81</u>	<u>0.87</u>	<u>0.93</u>	<u>0.96</u>	<u>0.99</u>	<u>1.00</u>	<u>1.00</u>
0.9	<u>0.69</u>	<u>0.71</u>	<u>0.75</u>	<u>0.77</u>	<u>0.80</u>	<u>0.83</u>	<u>0.87</u>	<u>0.90</u>	<u>0.95</u>	<u>1.00</u>	<u>1.00</u>
1.0	<u>0.78</u>	<u>0.79</u>	<u>0.79</u>	<u>0.81</u>	<u>0.83</u>	<u>0.84</u>	<u>0.85</u>	<u>0.87</u>	<u>0.92</u>	<u>0.99</u>	<u>1.00</u>

Table 1.2. Proportion d'échantillons pour lesquels $P(M_1/W') > 0.5$.
La loi a priori Beta1, n=10.

ρ	δ										
	0.0	1.0	1.5	2.0	2.5	3.0	3.5	4.0	5.0	10	20
$k = 5$											
0.0	0.02	0.07	0.15	0.30	0.45	<u>0.57</u>	<u>0.69</u>	<u>0.79</u>	<u>0.91</u>	<u>1.00</u>	<u>1.00</u>
0.1	0.03	0.09	0.19	0.33	0.49	<u>0.62</u>	<u>0.72</u>	<u>0.81</u>	<u>0.93</u>	<u>1.00</u>	<u>1.00</u>
0.3	0.06	0.15	0.26	0.42	<u>0.57</u>	<u>0.68</u>	<u>0.78</u>	<u>0.85</u>	<u>0.94</u>	<u>1.00</u>	<u>1.00</u>
0.5	0.14	0.25	0.37	<u>0.50</u>	<u>0.62</u>	<u>0.72</u>	<u>0.80</u>	<u>0.87</u>	<u>0.95</u>	<u>1.00</u>	<u>1.00</u>
0.7	0.29	0.38	0.47	<u>0.56</u>	<u>0.66</u>	<u>0.74</u>	<u>0.80</u>	<u>0.86</u>	<u>0.93</u>	<u>1.00</u>	<u>1.00</u>
0.9	<u>0.55</u>	<u>0.58</u>	<u>0.60</u>	<u>0.64</u>	<u>0.67</u>	<u>0.71</u>	<u>0.76</u>	<u>0.80</u>	<u>0.87</u>	<u>1.00</u>	<u>1.00</u>
1.0	<u>0.70</u>	<u>0.69</u>	<u>0.68</u>	<u>0.68</u>	<u>0.69</u>	<u>0.71</u>	<u>0.73</u>	<u>0.75</u>	<u>0.80</u>	<u>0.97</u>	<u>1.00</u>
$k = 2$											
0.0	0.01	0.05	0.11	0.22	0.35	0.48	<u>0.61</u>	<u>0.70</u>	<u>0.86</u>	<u>1.00</u>	<u>1.00</u>
0.1	0.02	0.06	0.14	0.28	0.40	<u>0.52</u>	<u>0.64</u>	<u>0.74</u>	<u>0.89</u>	<u>1.00</u>	<u>1.00</u>
0.3	0.05	0.12	0.22	0.35	0.49	<u>0.62</u>	<u>0.72</u>	<u>0.82</u>	<u>0.93</u>	<u>1.00</u>	<u>1.00</u>
0.5	0.12	0.22	0.33	0.47	<u>0.59</u>	<u>0.69</u>	<u>0.79</u>	<u>0.86</u>	<u>0.95</u>	<u>1.00</u>	<u>1.00</u>
0.7	0.30	0.37	0.47	<u>0.56</u>	<u>0.65</u>	<u>0.72</u>	<u>0.80</u>	<u>0.86</u>	<u>0.94</u>	<u>1.00</u>	<u>1.00</u>
0.9	<u>0.53</u>	<u>0.58</u>	<u>0.59</u>	<u>0.64</u>	<u>0.68</u>	<u>0.71</u>	<u>0.75</u>	<u>0.80</u>	<u>0.86</u>	<u>0.99</u>	<u>1.00</u>
1.0	<u>0.68</u>	<u>0.67</u>	<u>0.68</u>	<u>0.69</u>	<u>0.71</u>	<u>0.73</u>	<u>0.74</u>	<u>0.76</u>	<u>0.81</u>	<u>0.95</u>	<u>1.00</u>

Table 1.3. Proportion d'échantillons pour lesquels $P(M_1/W') > 0.5$.
La loi a priori triangulaire, n=100.

[illegible]

Table 1.4. Proportion d'échantillons pour lesquels $P(M_1/W') > 0.5$.
La loi a priori Beta1, n=100.

ρ	δ										
	0.0	1.0	1.5	2.0	2.5	3.0	3.5	4.0	5.0	10	20
$k = 50$											
0.0	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.02	<u>1.00</u>	<u>1.00</u>
0.1	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.03	<u>1.00</u>	<u>1.00</u>
0.3	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.16	<u>1.00</u>	<u>1.00</u>
0.5	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.04	0.43	<u>1.00</u>	<u>1.00</u>
0.7	0.00	0.00	0.00	0.00	0.01	0.04	0.12	0.29	<u>0.71</u>	<u>1.00</u>	<u>1.00</u>
0.9	0.20	0.25	0.30	0.35	0.42	<u>0.52</u>	<u>0.60</u>	<u>0.68</u>	<u>0.80</u>	<u>1.00</u>	<u>1.00</u>
1.0	<u>0.91</u>	<u>0.90</u>	<u>0.90</u>	<u>0.90</u>	<u>0.91</u>	<u>0.91</u>	<u>0.92</u>	<u>0.92</u>	<u>0.92</u>	<u>0.94</u>	<u>0.99</u>
$k = 5$											
0.0	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.04	<u>1.00</u>
0.1	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.06	<u>1.00</u>
0.3	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.21	<u>1.00</u>
0.5	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.49	<u>1.00</u>
0.7	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.01	<u>0.81</u>	<u>1.00</u>
0.9	0.22	0.23	0.24	0.27	0.29	0.32	0.35	0.39	0.49	<u>0.95</u>	<u>1.00</u>
1.0	<u>0.92</u>	<u>0.91</u>	<u>0.91</u>	<u>0.91</u>	<u>0.91</u>	<u>0.91</u>	<u>0.91</u>	<u>0.90</u>	<u>0.90</u>	<u>0.92</u>	<u>0.96</u>

Table 1.5. Proportion d'échantillons pour lesquels $P(M_1/W') > 0.5$.
La loi a priori Beta2, n=100.

ρ	δ										
	0.0	1.0	1.5	2.0	2.5	3.0	3.5	4.0	5.0	10	20
$k = 50$											
0.0	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	<u>0.88</u>	<u>1.00</u>
0.1	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	<u>0.93</u>	<u>1.00</u>
0.3	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	<u>0.99</u>	<u>1.00</u>
0.5	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.02	<u>1.00</u>	<u>1.00</u>
0.7	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.02	0.12	<u>1.00</u>	<u>1.00</u>
0.9	0.03	0.03	0.05	0.08	0.12	0.17	0.23	0.30	0.47	<u>0.98</u>	<u>1.00</u>
1.0	<u>0.78</u>	<u>0.78</u>	<u>0.78</u>	<u>0.78</u>	<u>0.78</u>	<u>0.79</u>	<u>0.78</u>	<u>0.78</u>	<u>0.79</u>	<u>0.82</u>	<u>0.93</u>
$k = 5$											
0.0	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.08	<u>1.00</u>
0.1	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.15	<u>1.00</u>
0.3	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.36	<u>1.00</u>
0.5	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	<u>0.63</u>	<u>1.00</u>
0.7	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	<u>0.86</u>	<u>1.00</u>
0.9	0.05	0.05	0.05	0.06	0.07	0.08	0.10	0.14	0.25	<u>0.94</u>	<u>1.00</u>
1.0	<u>0.80</u>	<u>0.79</u>	<u>0.79</u>	<u>0.78</u>	<u>0.78</u>	<u>0.79</u>	<u>0.78</u>	<u>0.79</u>	<u>0.78</u>	<u>0.84</u>	<u>0.92</u>

Table B3. Proportion d'échantillons pour lesquels $P(M_1/W') > 0.5$
Loi a priori Beta1, avec un AO.

ρ	Δ										
	0.0	1.0	1.5	2.0	2.5	3.0	3.5	4.0	5.0	10	20
	$k = 50$										
0.80	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
0.85	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.02	0.00
0.90	0.21	0.21	0.21	0.21	0.21	0.21	0.21	0.21	0.20	0.17	0.11
0.95	<u>0.59</u>	<u>0.58</u>	<u>0.57</u>	<u>0.57</u>	<u>0.58</u>	<u>0.58</u>	<u>0.58</u>	<u>0.58</u>	<u>0.59</u>	<u>0.60</u>	<u>0.68</u>
1.00	<u>0.91</u>	<u>0.90</u>	<u>0.91</u>	<u>0.91</u>	<u>0.91</u>	<u>0.90</u>	<u>0.90</u>	<u>0.90</u>	<u>0.91</u>	<u>0.92</u>	<u>0.99</u>
	$k = 5$										
0.80	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
0.85	0.02	0.02	0.02	0.02	0.02	0.03	0.03	0.03	0.02	0.01	0.00
0.90	0.19	0.19	0.19	0.19	0.19	0.19	0.19	0.18	0.18	0.16	0.09
0.95	<u>0.60</u>	<u>0.60</u>	<u>0.60</u>	<u>0.59</u>	<u>0.58</u>	<u>0.58</u>	<u>0.58</u>	<u>0.58</u>	<u>0.57</u>	<u>0.58</u>	<u>0.66</u>
1.00	<u>0.90</u>	<u>0.90</u>	<u>0.91</u>	<u>0.91</u>	<u>0.91</u>	<u>0.91</u>	<u>0.91</u>	<u>0.91</u>	<u>0.91</u>	<u>0.92</u>	<u>0.96</u>

Table B4. Proportion d'échantillons pour lesquels $P(M_1/W') > 0.5$
Loi a priori Beta1, avec trois AO.

ρ	Δ										
	0.0	1.0	1.5	2.0	2.5	3.0	3.5	4.0	5.0	10	20
0.80	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
0.85	0.04	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
0.90	0.26	0.16	0.09	0.06	0.04	0.04	0.03	0.02	0.02	0.00	0.00
0.95	<u>0.64</u>	<u>0.74</u>	<u>0.77</u>	<u>0.83</u>	<u>0.86</u>	<u>0.88</u>	<u>0.90</u>	<u>0.91</u>	<u>0.94</u>	<u>0.94</u>	<u>0.97</u>
1.00	<u>0.94</u>	<u>0.97</u>	<u>0.98</u>	<u>0.99</u>	<u>0.99</u>	<u>0.99</u>	<u>0.99</u>	<u>1.00</u>	<u>1.00</u>	<u>1.00</u>	<u>1.00</u>

Table B5. Proportion d'échantillons pour lesquels $P(M_1/W') > 0.5$
Loi a priori Beta2, avec un AO.

ρ	Δ										
	0.0	1.0	1.5	2.0	2.5	3.0	3.5	4.0	5.0	10	20
	$k = 50$										
0.80	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
0.85	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
0.90	0.04	0.04	0.04	0.04	0.04	0.04	0.03	0.03	0.03	0.01	0.00
0.95	0.26	0.26	0.26	0.25	0.25	0.24	0.24	0.23	0.23	0.19	0.09
1.00	<u>0.79</u>	<u>0.79</u>	<u>0.79</u>	<u>0.79</u>	<u>0.79</u>	<u>0.79</u>	<u>0.78</u>	<u>0.78</u>	<u>0.78</u>	<u>0.81</u>	<u>0.90</u>
	$k = 5$										
0.80	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
0.85	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
0.90	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.01	0.01	0.00
0.95	0.24	0.24	0.23	0.23	0.23	0.23	0.22	0.21	0.21	0.16	0.09
1.00	<u>0.79</u>	<u>0.79</u>	<u>0.79</u>	<u>0.79</u>	<u>0.79</u>	<u>0.79</u>	<u>0.79</u>	<u>0.79</u>	<u>0.80</u>	<u>0.82</u>	<u>0.92</u>

Table B6. Proportion d'échantillons pour lesquels $P(M_1/W') > 0.5$
Loi a priori Beta2, avec trois AO.

ρ	Δ										
	0.0	1.0	1.5	2.0	2.5	3.0	3.5	4.0	5.0	10	20
0.80	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
0.85	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
0.90	0.05	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
0.95	0.30	0.15	0.09	0.06	0.03	0.03	0.02	0.02	0.02	0.00	0.00
1.00	<u>0.85</u>	<u>0.91</u>	<u>0.95</u>	<u>0.97</u>	<u>0.98</u>	<u>0.97</u>	<u>0.98</u>	<u>0.99</u>	<u>1.00</u>	<u>1.00</u>	<u>1.00</u>