

MINISTÈRE DE L'ENSEIGNEMENT ET DE LA RECHERCHE  
SCIENTIFIQUE  
UNIVERSITÉ MOULOUD MAMMERRI, TIZI-OUZOU  
FACULTÉ DES SCIENCES  
DÉPARTEMENT DE MATHÉMATIQUES



**MÉMOIRE DE MASTER**

Filière : Mathématique

Spécialité : Probabilités et statistique

Présenté par

Mlle Sidhoumi Thafsouth

Sujet

---

**Analyse Bayésienne d'un modèle de régression linéaire  
sous contamination AO**

---

Devant le jury composé de :

BOUALAM Karima :	MCB	UMMTO	Présidente
AIT MOHAMMED Noura :	MAA	UMMTO	Promotrice
MERABET Dalila :	MCB	UMMTO	Examinatrice

Année universitaire : 2022/2023

# Remerciements

Je remercie Dieu tout puissant de qui par la grâce les bonnes actions sont accomplies de m'avoir donnée la santé, la persévérance, et la volonté d'entamer et de terminer ce mémoire.

Tout d'abord, je remercie ma promotrice **Madame Ait Mohammed** épouse **Graiche** sans qui ce travail n'aurait pas pu voir le jour, je la remercie pour la qualité de son encadrement, sa patience, et sa disponibilité durant la préparation de ce mémoire.

J'aimerais aussi remercier **Madame Boualam** d'avoir présidé le jury et **Madame Merabet** d'avoir examiné ce travail.

Je remercie mes parents, qui ont toujours été là pour moi, et qui se sont sacrifiés n'épargnant ni santé ni effort pour faire de moi ce que je suis aujourd'hui. Je remercie aussi ma soeur **Ouiza** pour ses encouragements et son aide précieuse à l'élaboration de ce mémoire.

Je souhaite exprimer ma reconnaissance envers toutes les personnes qui ont joué un rôle, qu'il soit direct ou indirect, dans la réussite de ce travail.

# Dédicace

*Avec l'expression de ma reconnaissance, je dédie ce modeste travail à mes chers parents et ma chère soeur, qui m'ont toujours poussée et motivée dans mes études. Ce projet représente donc l'aboutissement du soutien et des encouragements qu'ils m'ont prodigués tout au long de ma scolarité. Qu'ils en soient remerciés par cette trop modeste dédicace.*

# Table des matières

<b>Table des matières</b>	<b>2</b>
<b>Notation</b>	<b>4</b>
<b>Introduction générale</b>	<b>6</b>
<b>1 Notions générales et définitions</b>	<b>9</b>
1.1 Introduction . . . . .	9
1.2 Inférence Bayésienne . . . . .	10
1.2.1 Modèles Bayésiens . . . . .	10
1.2.1.1 Théorème de Bayes . . . . .	11
1.2.2 Loi a priori . . . . .	11
1.2.2.1 Lois a priori informatives . . . . .	12
1.2.2.2 Lois a priori non informatives . . . . .	16
1.2.3 Estimateur de Bayes . . . . .	17
1.2.4 Fonctions coût usuelles . . . . .	19
1.3 Méthode de calcul Bayésien . . . . .	21
1.3.1 Méthode de Monte Carlo par chaîne de Markov . . . . .	21
1.3.1.1 Algorithme de Metropolis-Hastings . . . . .	21
1.3.1.2 L'échantillonnage de Gibbs . . . . .	22
1.4 Modèles autorégressifs (AR) . . . . .	24

<b>2</b>	<b>Inférence Bayésienne d'un modèle de régression linéaire sous contamination</b>	<b>26</b>
2.1	Introduction . . . . .	26
2.2	Régression linéaire simple . . . . .	26
2.2.1	Hypothèses du modèle . . . . .	27
2.3	Estimation des paramètres d'un modèle de régression linéaire quand les erreurs sont dépendantes . . . . .	28
2.3.1	Le modèle . . . . .	28
2.3.2	Lois a posteriori conditionnelles de $\beta$ et $\alpha$ . . . . .	29
2.3.2.1	Loi a posteriori conditionnelle de $\beta$ . . . . .	31
2.3.2.2	Loi a posteriori conditionnelle de $\alpha$ . . . . .	32
2.4	Estimation des paramètres d'un modèle de régression linéaire quand les erreurs sont dépendantes ( $\epsilon_t \sim AR(1)$ ) sous contamination AO . . . . .	34
2.4.1	Présentation du modèle AO . . . . .	34
2.4.2	Lois a posteriori conditionnelles de $\beta, \alpha, \Delta$ . . . . .	35
2.4.2.1	Loi a posteriori de $\beta$ sachant $\alpha$ et $\Delta$ . . . . .	36
2.4.2.2	Loi a posteriori de $\alpha$ sachant $\Delta$ et $\beta$ . . . . .	38
2.4.2.3	Loi a posteriori de $\Delta$ sachant $\alpha$ et $\beta$ . . . . .	39
2.5	Densité prédictive d'une observation future sous contamination AO . . . . .	41
2.5.1	Lois a posteriori conditionnelles de $y_{n+1}, \beta, \alpha, \Delta$ . . . . .	42
2.5.1.1	Loi a posteriori conditionnelle de $y_{n+1}$ . . . . .	42
2.5.1.2	Loi a posteriori conditionnelle de $\alpha$ . . . . .	44
2.5.1.3	Loi a posteriori conditionnelle de $\beta$ . . . . .	45
2.5.1.4	Loi a posteriori conditionnelle de $\Delta$ . . . . .	47
<b>3</b>	<b>Étude de simulation</b>	<b>49</b>
3.1	Présentation du modèle . . . . .	50

3.2	Application de l'échantillonneur de Gibbs pour l'estimation des paramètres du modèle sous contamination AO . . . . .	51
3.3	Application de l'échantillonneur de Gibbs pour la prédiction d'une valeur future sous contamination AO . . . . .	54
	<b>Conclusion générale</b>	<b>55</b>
	<b>Résumé</b>	<b>56</b>
	<b>Bibliographie</b>	<b>57</b>

# Notations

$\theta, \alpha, \beta, \sigma, \Delta$	Paramètres inconnus
$\hat{\alpha}, \hat{\beta}, \hat{\sigma}, \hat{\Delta}$	Estimateurs de $\theta, \alpha, \beta, \sigma, \Delta$
$\Theta$	Ensemble des paramètres
$s^2$	variance empirique
iid	indépendantes et identiquement distribuées
$\mathbb{N}$	Ensemble des entiers naturels
$C$	Constante de normalisation
$\mathbb{Z}$	Ensemble des nombres entiers relatifs
$\mathbb{R}$	Ensemble des réels
$\pi$	Distribution de probabilité
$\mathbb{R}^p$	Ensemble des vecteurs réels à p dimensions
$\mathbb{E}$	Espérance mathématique

$\mathcal{F}$	Famille des lois sur $\Theta$
$MCMC$	Markov chain Monte Carlo
$AR(p)$	processus autorégressif d'ordre $p$
$\mathcal{N}(0, \sigma^2)$	Loi Normal centrée
$\propto$	proportionnel
$X$	l'espace des observations
$\Theta$	l'espace des états de la nature (l'espace des paramètres dans le cas d'un problème statistique)
$A$	l'espace des actions ou décisions dont les éléments sont des images de l'observation par une application $\delta$
$D$	l'ensemble des règles de décisions $\delta$ applications de $X$ dans $A$ (les estimateurs possibles).
$\delta$	règle de décisions (une statistique (i.e. fonction de observations) dans le cas d'un problème statistique)
v.a	variable aléatoire
$\mathbb{E}^\pi$	espérance associée à la loi $\pi$
$\mathbb{E}_\theta$	espérance par rapport à une distribution paramétrisée par le paramètre $\theta$
rmse	Root Mean Square Error (Erreur Quadratique Moyenne (EQM))
std	standard deviation (l'écart-type)

# Introduction générale

” Statistical thinking will one day be as necessary for efficient citizenship as the ability to read and write.”

H.G.Wells

L’analyse bayésienne est une approche statistique qui permet d’inférer des conclusions sur des paramètres inconnus à partir de données observées en combinant des informations a priori et des données fondées sur l’expérience et l’observation. Lorsqu’elle est appliquée à un modèle de régression linéaire simple, elle offre plusieurs avantages et apporte une valeur significative à l’analyse statistique.

Dans l’analyse des données statistiques, notre intérêt se porte sur l’estimation de divers paramètres selon le modèle spécifique étudié. Plus précisément, nous allons examiner le modèle de régression linéaire avec erreurs dépendantes  $\epsilon_t \sim AR(1)$  sous contamination AO. Dans ce contexte, certaines références pertinentes sont les travaux de [Zellner et Tiao \(1964\)](#), [Parent et Bernier \(2007\)](#), [John et al \(1996\)](#), ainsi que [Girard et Parent \(2000\)](#).

La contamination AO (Additive Outlier en anglais) est un concept statistique utilisé pour décrire des observations qui diffèrent nettement du reste de l’échantillon. Ces observations, appelées valeurs aberrantes, se distinguent par leur

---

écart significatif par rapport aux autres valeurs de l'échantillon. La contamination AO peut se produire dans divers domaines de l'analyse de données, tels que la finance, l'économie, les sciences sociales, l'environnement, etc.

Une valeur aberrante peut résulter de divers facteurs, tels que des erreurs de mesure, des erreurs de saisie de données, des phénomènes rares ou inhabituels, ou des conditions expérimentales anormales.

Les valeurs aberrantes peuvent fausser les estimations des paramètres statistiques, influencer les modèles de prédiction et introduire des biais dans les analyses. Par conséquent, il est important d'identifier et de comprendre les valeurs aberrantes afin de prendre des décisions éclairées quant à leur traitement.

Il convient de noter que l'analyse du modèle de régression linéaire avec erreurs dépendantes ne peut être réalisée de manière explicite. Cependant, ce modèle présente une structure conditionnelle appropriée qui permet d'effectuer une inférence bayésienne en utilisant l'échantillonnage de Gibbs.

La méthode bayésienne repose sur l'estimation des paramètres à partir d'une distribution appelée loi a posteriori, qui combine l'information contenue dans les données (via la vraisemblance) avec une information a priori (telle que des avis d'expert ou des contraintes physiques). Cette approche est largement utilisée dans l'inférence statistique. Des exemples de références traitant de cette méthode sont les ouvrages de [Robert \(2006,2013\)](#) qui décrivent en détail les fondements et les applications de l'inférence bayésienne.

L'objectif de ce mémoire est d'une part se familiariser avec l'approche bayésienne et d'autre part montrer l'efficacité des méthodes MCMC plus particulièrement l'échantillonnage de Gibbs pour l'estimation bayésienne des

---

paramètres d'un modèle de régression linéaire avec erreurs autocorrélées sous contamination AO.

Dans ce mémoire, nous abordons trois chapitres essentiels.

Le premier chapitre fournit un aperçu des concepts et des rappels généraux nécessaires à la compréhension du reste du document. Nous discutons des lois a priori, à la fois informatives et non informatives, ainsi que des méthodes de calcul bayésien. De plus, nous présentons quelques notions sur les séries chronologiques, notamment les modèles autorégressifs et l'opérateur retard.

Nous consacrons le deuxième chapitre à l'inférence bayésienne d'un modèle de régression sous contamination. Nous commençons par définir le modèle de régression linéaire simple et discutons de l'estimation des paramètres dans le cas où les erreurs sont dépendantes en utilisant l'échantillonnage de Gibbs. Ensuite, nous déterminons les différentes lois a posteriori conditionnelles des paramètres. Par la suite, nous introduisons une contamination dans notre modèle, en utilisant par exemple une contamination AO. Nous déterminons alors les lois a posteriori conditionnelles des différents paramètres dans ce modèle de contamination. Enfin, nous abordons la densité prédictive d'une observation future sous contamination AO et les lois a posteriori conditionnelles des différents paramètres associées à cette prédiction.

Dans le dernier chapitre, nous donnons les résultats de simulations pour justifier l'efficacité des résultats obtenus dans le chapitre précédent.

Ce mémoire se termine par une conclusion générale et d'un résumé récapitulatif.

# Chapitre 1

## Notions générales et définitions

### 1.1 Introduction

L'analyse Bayésienne nommée d'après Thomas Bayes et développée par Simon de Laplace est un outil analytique puissant pour la modélisation statistique, elle répond aux questions de recherche sur des paramètres inconnus à l'aide d'énoncés de probabilité. La première étape de cette analyse consiste à étudier la situation et d'identifier une incertitude portée sur un paramètre inconnu  $\theta$ , cette incertitude est modélisée sous la forme d'une distribution dite a priori, qui donne une information sur  $\theta$  pris comme étant une variable aléatoire.

Le schéma ci-dessous résume la démarche bayésienne :

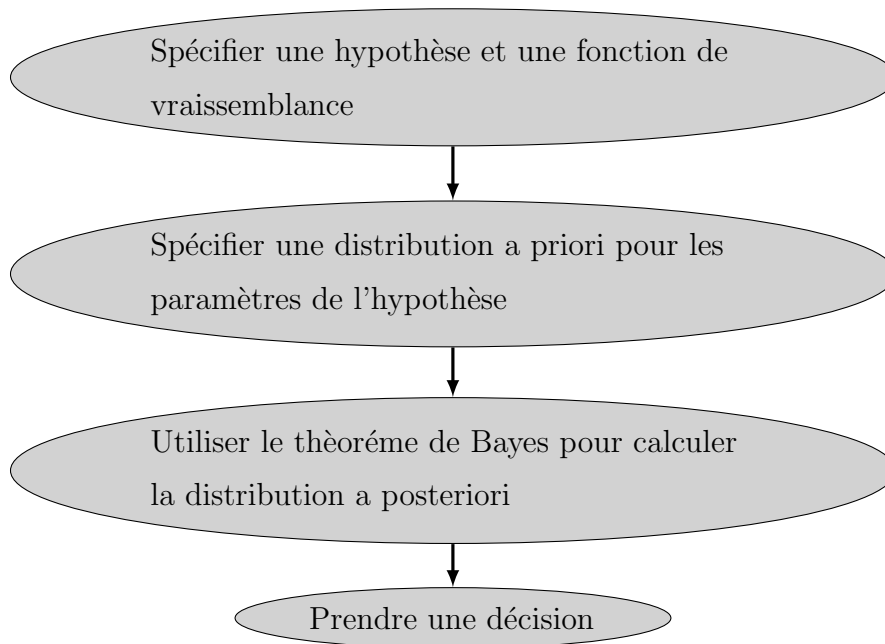


FIGURE 1.1 – La démarche bayésienne

## 1.2 Inférence Bayésienne

En se plaçant dans un cadre de modèle statistique paramétrique avec  $X = (x_1, x_2, \dots, x_n)$  un échantillon de données, il est primordial de faire de l'inférence (estimations, tests d'hypothèses,...) et de la prévision au sujet d'un ou plusieurs paramètres d'intérêts  $\theta$ . Pour cela, on distingue deux approches dont la différence réside dans la façon de traiter le paramètre  $\theta$ . Selon la première approche qui est l'approche classique (fréquentielle),  $\theta$  prend une valeur inconnue fixe, pour la deuxième qui est la méthode Bayésienne,  $\theta$  est vu comme aléatoire.

Pour une compréhension plus détaillée des concepts et des méthodes de la statistique Bayésienne, nous recommandons [Bernardo et Smith \(2009\)](#), [Carlin et Louis \(1997\)](#), [Gelman et al. \(1995\)](#) et [Robert \(2006\)](#).

### 1.2.1 Modèles Bayésiens

**Définition 1.1.** *On appelle modèle statistique Bayésien la donnée d'un modèle statistique paramétrique  $(X, A, P_\theta, \theta \in \Theta)$  avec  $f(x|\theta)$  densité de  $\pi(\theta)$  et d'une loi  $\pi(\theta)$  sur le paramètre.*

Définissons les lois de probabilités intervenants dans l'analyse Bayésienne :

- **La vraisemblance**

C'est la loi des observations ou encore la loi conditionnelle de  $X$  sachant  $\theta$ , notée par  $f(x|\theta)$ , elle est donnée par :

$$f(x|\theta) = \prod_{i=1}^n f(x_i|\theta)$$

- **Loi a priori**

Notée par  $\pi(\theta)$ , désigne la densité du paramètre inconnu  $\theta$ , elle porte l'information sur ce dernier. L'appellation a priori exprime le fait qu'elle a été établie préalablement à l'observation des données  $x$ .

- **Loi jointe du couple**  $(\theta, x)$

Généralement notée par  $f(\theta, x)$ , sa formule est donnée par :

$$f(\theta, x) = f(x|\theta)\pi(\theta)$$

**1.2.1.1 Théorème de Bayes**

Soit  $X = (x_1, x_2, \dots, x_n)$  un échantillon de  $n$  observations de densité  $f(x | \theta)$  et de distribution a priori  $\pi(\theta)$ , la démarche de l'analyse Bayésienne conduit au calcul d'une loi a posteriori  $\pi(\theta | x)$ , actualisation de la loi a priori  $\pi(\theta)$  au vu de l'observation.

Ce calcul repose sur la version continue du théorème de Bayes :

$$\pi(\theta | x) = \frac{f(x | \theta)\pi(\theta)}{f(x)} \quad (1.1)$$

avec  $f(x | \theta)$  : la loi de l'observation (vraisemblance) et  $f(x)$  : la loi marginale (prédictive) :

$$f(x) = \int_{\Theta} f(x | \theta)\pi(\theta)d\theta$$

**Remarque 1.1.** *La loi a posteriori représentée dans l'équation (1.1) est la loi de  $\theta$  sachant  $x$ . En vertu de la loi de Bayes on peut l'écrire de la manière suivante :*

$$\begin{aligned} \pi(\theta | x) &= \frac{f(x | \theta)\pi(\theta)}{f(x)} \\ &= \frac{f(x | \theta)\pi(\theta)}{\int_{\Theta} f(x | \theta)\pi(\theta)d\theta} \end{aligned} \quad (1.2)$$

**1.2.2 Loi a priori**

Le choix de la loi a priori est le point le plus contestable de l'analyse Bayésienne, car une fois connue, l'inférence sera conduite d'une manière quasi-mécanique. Sa détermination est donc l'étape la plus importante de cette inférence. Ce choix, peut être basé sur différents points de vues tels que : les expériences du passé ou

une intuition du statisticien, la faisabilité des calculs. La loi a priori de  $\theta$ , notée  $\pi(\cdot)$  modélise l'information disponible sur  $\theta$  avant d'observer les données.

On va chercher non pas "la meilleure valeur" de ce paramètre comme dans le maximum de vraisemblance, mais on va estimer la distribution de probabilité de ce paramètre.

Dans la suite de ce document, nous allons voir que les lois a priori sont divisés en deux grandes parties qui sont : *l'approche des objectivistes* où, on suppose une probabilité a priori dite *distribution a priori non informative*, du fait que l'observateur n'a à l'origine rien à sa disposition mis à part son outil de mesure, ainsi que *"l'approche des subjectivistes"* où on suppose la probabilité a priori dite *la distribution a priori informative* qui permet d'inclure des connaissances préexistantes à la mesure.

### 1.2.2.1 Lois a priori informatives

Une loi a priori informative est une distribution de probabilité sélectionnée de manière à inclure des informations antérieures ou des connaissances d'experts sur les valeurs possibles des paramètres d'un modèle statistique. Elle peut fortement influencer l'estimation des paramètres en fournissant des hypothèses ou des contraintes supplémentaires. (Voir : [Robert\(2006,2010\)](#), [Casella \(1992\)](#), [Freedman et Spiegelhalter \(1983\)](#) et [Berry \(1985\)](#) )

### Loi a priori conjuguées

Le choix d'une loi a priori conjuguée bien qu'il soit défendable est toujours un choix particulier et influence donc, dans une certaine mesure, l'inférence résultante. Il est souvent choisi tel que sa conjugaison avec la vraisemblance produise une loi a posteriori de la même famille.

**Définition 1.2.** Une famille  $\mathcal{F}$  de distribution sur  $\Theta$  est dite conjuguée pour la loi  $f(X | \theta)$  si  $\forall \pi \in \mathcal{F}$ , la distribution a posteriori  $\pi(\cdot | X)$  appartient également à  $\mathcal{F}$ . L'avantage des familles conjuguées est avant tout de simplifier les calculs.

**Remarque 1.2.** Une loi conjuguée peut être déterminée en considérant la forme de la vraisemblance  $f(X | \theta)$  et en prenant une loi a priori de la même forme. Les lois a posteriori obtenues par ce procédé sont dites conjuguées naturelles.

**Exemple 1.1.** le tableau suivant montre des exemples de lois a priori conjuguées

$f(x \theta)$	$\pi(\theta)$	$\pi(\theta x)$
Normale $\mathcal{N}(\theta, \sigma^2)$	Normale $\mathcal{N}(\mu, \tau^2)$	$\mathcal{N}(\rho(\sigma^2\mu + \tau^2x), \rho\sigma^2\tau^2)$
Poisson $\mathcal{P}(\theta)$	Gamma $\Gamma(\alpha, \beta)$	$\Gamma(\alpha + x, \beta + 1)$
Gamma $\mathcal{P}(\theta)$	Gamma $\Gamma(\alpha, \beta)$	$\Gamma(\alpha + v, \beta + x)$
Binomiale $\mathcal{B}(n, \theta)$	Beta $\mathcal{B}e(\alpha, \beta)$	$\mathcal{B}e(\alpha + m, \beta + x)$
Multinomiale $\mathcal{M}_k(\theta_1, \dots, \theta_k)$	Dirichlet $\mathcal{D}(\alpha_1, \dots, \alpha_k)$	$\mathcal{D}(\alpha_1 + x_1, \dots, \alpha_k + x_k)$
Normale $\mathcal{N}(\mu, \frac{1}{\theta})$	Gamma $\Gamma(\alpha, \beta)$	$\Gamma(\alpha + 0.5, \beta + \frac{(\mu-x)^2}{2})$

TABLE 1.1 – Lois a priori conjuguées usuelles

En général, les lois a priori conjuguées sont reliées à un type particulier de lois d'échantillonnage qui permet toujours leur obtention, et comme nous le verrons ci-dessous, ces lois constituent des familles appelées familles exponentielles.

**Exemple 1.2.** On considère  $X$  comme étant une binomiale négative  $Neg(n, \theta)$  définie par :

$$P[X = x|\theta] = C_{n+x-1}^x \theta^x (1 - \theta)^n \quad 0 < \theta < 1, x \in \mathbb{N}, n \in \mathbb{N}^* \quad (1.3)$$

Une loi conjuguée naturelle serait une loi bêta car  $P(X = x|\theta) \propto \theta^x(1 - \theta^n)$  qui ressemble à une loi bêta.

**Exemple 1.3.** Soit  $X$  une v.a de loi de Pareto de paramètres  $\theta$  et  $a$ ,  $a > 0$ ,  $\theta > 0$

$$\begin{aligned} f(x|\theta, a) &= \frac{\theta a^\theta}{x^{\theta+1}}, \quad x \geq a \\ &= \theta e^{\theta \ln(a)} e^{-(\theta+1)\ln(x)} \\ &= \theta e^{\theta(\ln(a)-\ln(x))} e^{-\ln(x)} \\ &= \theta e^{\theta \ln(\frac{a}{x})} e^{-\ln(x)} \\ &\propto \theta e^{\theta \ln(\frac{a}{x})} \end{aligned}$$

On peut donc prendre une loi a priori Gamma

### Familles exponentielles

**Définition 1.3.** Toute famille à  $s$  paramètres de loi de distribution  $\{P_\theta\}$  dont la densité de probabilité est donnée de la forme suivante :

$$f(x|\theta) = h(x) \cdot \exp\left(\sum_{i=1}^s \eta_i T_i(x) - \beta(\theta)\right) \quad (1.4)$$

est dite famille exponentielle.

avec  $\eta_i$ ,  $\beta(\cdot)$  sont des fonction de  $\theta$ .

$T_i(\cdot)$  et  $h(\cdot)$  sont des statistiques (fonction de  $X$ ).

**Définition 1.4.** La forme canonique de cette densité est donnée en reparamétrisant  $\eta_i(\theta)$  par  $\theta_i$  par l'équation ci-dessous :

$$f(x|\theta) = h(x) \exp\left(\sum_{i=1}^s \theta_i T_i(x) - \beta(\theta)\right) \quad (1.5)$$

**Proposition 1.1.** *Robert (2006)*

Une famille conjuguée pour  $f(x|\theta)$  est donnée par

$$\pi(\theta|\mu, \lambda) = K(\mu, \lambda)\exp(\theta\mu - \lambda A(\theta)) \quad (1.6)$$

où  $K(\mu, \lambda)$  est la constante de normalisation de la densité. La loi a posteriori correspondante est donnée comme suit

$$\pi(\theta | x) \propto \exp\left((\mu + x)\theta - (\lambda + 1)A(\theta)\right) \quad (1.7)$$

**Exemple 1.4.** La loi gamma  $(\alpha, \beta)$

$$f(x|\alpha, \beta) = \frac{\beta^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-\beta x} \quad \theta = (\alpha, \beta)$$

$$\begin{aligned} f(x | \alpha, \beta) &= \exp\left(\text{Ln}\frac{\beta^\alpha}{\Gamma(\alpha)} + (\alpha - 1)\text{Ln}(x) - \beta x\right) \\ &= \exp\left((\alpha - 1)\text{ln}(x) - \beta x + \text{Ln}\frac{\beta^\alpha}{\Gamma(\alpha)}\right) \end{aligned}$$

avec  $s = 2$ ,  $\eta_i(\theta) = (\alpha - 1)$ ,  $T_1(x) = \text{ln}(x)$ ,  $\eta_2(\theta) = \beta$ ,  $T_2(x) = x$ ,  
 $-\beta(\theta) = \text{Ln}(\frac{\beta^\alpha}{\Gamma(\alpha)})$ ,  $h(x) = \mathbb{1}$

### L'écriture canonique

$$\begin{aligned} f(x | \alpha, \beta) &= \exp\left(\alpha \text{Ln}(x) - \beta x + \text{Ln}\frac{\beta^\alpha}{\Gamma(\alpha)} - \text{Ln}(x)\right) \\ &= \exp\left(\alpha \text{Ln}(x) - \beta x + \text{Ln}\frac{\beta^\alpha}{\Gamma(\alpha)}\right) e^{-\text{Ln}x} = \frac{1}{x} \\ &= \exp\left(\alpha \text{Ln}(x) - \beta x + \text{Ln}\frac{\beta^\alpha}{\Gamma(\alpha)}\right) = \frac{1}{x} \end{aligned}$$

- La loi a priori conditionnel

$$\pi(\theta | \lambda, \mu) = K(\mu, y)\exp\left(\alpha\mu_1 + \beta\mu_2 + \lambda \frac{\text{Ln}\beta^\alpha}{\Gamma(\alpha)}\right)$$

- La loi a posteriori conditionnelle aux paramètres  $x, \lambda, \mu$

$$\pi(\theta | x, \lambda, \mu) \propto \exp\left(\alpha(\mu_1 + x) + \beta(\mu_2 + x) + (\lambda + 1) \frac{\text{Ln}\beta^\alpha}{\Gamma(\alpha)}\right)$$

### 1.2.2.2 Lois a priori non informatives

**Définition 1.5.** *Une loi non informative est une loi qui porte une information sur le paramètre à estimer dont le poids dans l'inférence est réduit.*

#### Loi a priori uniforme

La loi a priori non informative la plus simple est la loi uniforme. Ce choix repose sur l'équiprobabilité de  $\theta$ , par exemple si l'on dispose d'un paramètre inconnu  $\theta$  qui peut prendre n'importe quelle valeur réelle sur  $[a, b]$ , la loi uniforme assignera une probabilité à chaque valeur de  $\theta$  dans l'intervalle  $[a, b]$

$$p(\theta) = \frac{1}{(b-a)} \quad , \quad a \leq \theta \leq b$$

#### Les lois a priori de Laplace

Laplace fut le premier à utiliser des techniques non-informatives dans le cas d'absence d'informations, il utilise alors la loi uniforme qui est l'une des lois les plus simples et les plus utilisées parmi les lois a priori pour l'approche non-informative. En effet, ce choix repose sur l'équiprobabilité des valeurs du paramètre  $\theta$  dans son domaine de définition. Supposons que  $\Theta$  est un ensemble de taille  $k$  alors :

$$\pi(\theta) = \frac{1}{k}.$$

#### La loi a priori de Jeffreys

Proposée par Sir Harold Jeffreys en 1946, les lois non informatives de Jeffreys sont fondées sur l'information de Fisher donnée par :

$$I_x(\theta) = \mathbb{E}_\theta \left[ \left( \frac{\partial \text{Log} f(x|\theta)}{\partial \theta} \right)^2 \right]$$

qui est, dans le cas multidimensionnel sous condition que le domaine de  $X$  soit indépendant de  $\theta$ , égal à

$$I(\theta) = -\mathbb{E}_\theta \left[ \left( \frac{\partial^2 \text{Log}f(x|\theta)}{\partial^2 \theta} \right) \right]$$

La loi a priori de Jeffreys est

$$\pi_j(\theta) \propto C \sqrt{I_x(\theta)}$$

avec  $C$  une constante de normalisation.

Dans le cas multidimensionnel, le paramètre  $\theta$  est un vecteur  $(\theta_1, \theta_2, \dots, \theta_q)$ , d'où l'estimation des lois de la forme

$$\pi(\theta) = \left[ \det I(\theta) \right]^2$$

Où  $I(\theta)$  est la matrice de l'information de Fisher dont les éléments sont donnés par :

$$I_{i,j}(\theta) = -\mathbb{E}_\theta \left[ \frac{\partial^2 \log f(x|\theta)}{\partial \theta_i \partial \theta_j} \right] \quad 1 \leq i, j \leq q$$

### 1.2.3 Estimateur de Bayes

l'estimateur de Bayes combine des informations probabilistes et des données pour produire une estimation précise et robuste. Cette estimation est basée sur une probabilité a priori connue et prend en compte les informations supplémentaires à mesure que les données supplémentaires sont collectées.

On remarque que pour la prise d'une décision, le choix d'un estimateur, va engendrer un coût que l'on va quantifier à l'aide de la fonction de perte. Pratiquement, on cherche une décision qui minimise en moyenne la fonction de coût.

Soit  $L(\theta, \delta)$  une fonction de coût,  $\pi$  une loi a priori (une mesure), afin de trouver l'estimateur de Bayes  $\delta^\pi(x)$ , on applique la règle suivante

$$\delta^\pi(x) = \min_{\delta} \mathbb{E}^\pi[L(\theta, \delta)/x]$$

L'estimateur  $\delta^\pi$  sera déterminé analytiquement ou numériquement selon la complexité de la fonction perte et sa nature.

Généralement, les caractéristiques usuelles d'une distribution (moyenne, médiane,...,ect) correspondent aux solutions associées à des coûts classiques.

Par exemple l'estimateur de Bayes associé au coût quadratique est la moyenne a posteriori.

## Quelques estimateurs de Bayes

[Robert \(2006\)](#)

Quelques estimateurs de Bayes du paramètre  $\theta$  sous coût quadratique pour les lois *a priori conjuguées* des familles exponentielle usuelles avec  $\bar{x}$  est la moyenne des données observées.

Loi de x	loi conjuguée	moyenne a posteriori
Normale $\mathcal{N}(\theta, \sigma^2)$	Normale $\mathcal{N}(\mu, \tau^2)$	$\frac{\mu\sigma^2 + \tau^2\bar{x}}{\sigma^2}$
Poisson $\mathcal{P}(\theta)$	Gamma $\Gamma(\alpha, \beta)$	$\frac{\alpha + \bar{x}}{\beta + 1}$
Gamma $\mathcal{P}(\theta)$	Gamma $\Gamma(\alpha, \beta)$	$\frac{\alpha + v}{\beta + \bar{x}}$
Binomiale $\mathcal{B}(n, \theta)$	Beta $\mathcal{B}e(\alpha, \beta)$	$\frac{\alpha + x}{\alpha + \beta + n}$
Multinomiale $\mathcal{M}_k(\theta_1, \dots, \theta_k)$	Dirichlet $\mathcal{D}(\alpha_1, \dots, \alpha_k)$	$\frac{\alpha_i + x_i}{(\sum_j \alpha_j) + n}$
Normale $\mathcal{N}(\mu, \frac{1}{\theta})$	Gamma $\Gamma(\alpha, \beta)$	$\left(\frac{\alpha + 1}{\beta + (\mu - x)}\right)^2$

TABLE 1.2 – Quelques estimateurs de Bayes

### 1.2.4 Fonctions coût usuelles

Les fonctions coût usuelles les plus courantes sont *la fonction coût quadratique* et *la fonction coût absolue*.

**Définition 1.6.** On appelle *fonction coût* (ou *fonction perte*), toute fonction  $L$ ,

$$\begin{aligned} L : \Theta * A &\longmapsto R^+ \\ (\theta, \delta(x)) &\longmapsto L(\theta, \delta(x)) \end{aligned}$$

$L(\theta, \delta(x))$  évalue le coût associé à la décision  $a = \delta(x)$ .

Quand le paramètre vaut  $\theta$ , nous allons définir trois fonctions coût usuelles :

#### Le coût quadratique

**Définition 1.7.** La *fonction coût quadratique* est définie par

$$L(\theta, \delta(x)) = (\theta - \delta(x))^2$$

**Proposition 1.2.** *Robert(2006)*

L'estimateur de Bayes  $\delta^\pi$  associé à la loi a priori  $\pi$  et au coût quadratique est la

moyenne a posteriori

$$\delta^\pi(x) = E^\pi(\theta|x) = \frac{\int_{\Theta} \theta f(\theta|x)\pi(\theta)d\theta}{\int_{\Theta} f(\theta|x)\pi(\theta)d\theta}$$

**Corollaire 1.1.** *Robert(2006)*

Quand  $\theta \in R^p$ , l'estimateur de Bayes  $\delta^\pi$  associé à  $\pi$  et au coût quadratique,

$$L(\theta, \delta(x)) = (\theta - \delta)^t Q (\theta - \delta)$$

est la moyenne a posteriori,  $\delta^\pi(y) = \mathbb{E}^\pi[\theta|x]$ , pour toute matrice  $Q_{p \times p}$  symétrique définie positive.

### La fonction coût absolue

La fonction coût absolue s'écrit comme suit :

$$L(\theta, d) = |\theta - d|$$

généralement définie par morceaux

$$L(\theta, d) = \begin{cases} k_1(\theta - d) & \text{si } \theta > d \\ k_2(d - \theta) & \text{sinon} \end{cases}$$

**Proposition 1.3.** *Robert(2006)*

Un estimateur de Bayes associé à la loi a priori  $\pi$  et au coût absolue, est un fractile d'ordre  $k_2/(k_1 + k_2)$  de  $\pi(\theta|x)$

En particulier, si  $k_1 = k_2$ , dans le cas du coût absolu, l'estimateur de Bayes est la médiane a posteriori, qui est l'estimateur obtenu par Laplace.

## 1.3 Méthode de calcul Bayésien

### 1.3.1 Méthode de Monte Carlo par chaîne de Markov

Dans cette section, nous verrons des méthodes qui nous permettront d'approximer la génération de variables aléatoires d'une loi a posteriori  $\pi(\theta|x)$ .

L'une des premières versions de l'algorithme MCMC a été donnée en 1953 par Metropolis et al. dans un contexte de physique statistique et a été généralisé par Hastings en 1977. L'idée principale derrière ces méthodes consiste à construire une chaîne de Markov  $\{\theta_t\}$  de longueur significative, dont les échantillons convergent asymptotiquement vers la distribution cible souhaitée  $\pi(\theta|x)$ .

Dans la suite de ce document, nous allons présenter deux de ces méthodes conçues pour les chaînes de Markov de lois stationnaires données.

#### 1.3.1.1 Algorithme de Metropolis-Hastings

On peut décrire l'algorithme de Metropolis-Hastings de la façon suivante. Pour une densité donnée  $\pi(\theta)$ , connue à un facteur de normalisation près, et une densité conditionnelle  $q(\theta'|\theta)$ , la chaîne  $(\theta(m))_m$  est générée par l'algorithme comme suit :

- **itération 0** : Initialiser avec une valeur arbitraire  $\theta$
  
- **Itération m** : Mettre à jour  $\theta^{(m)}$  par  $\theta^{(m+1)}$  ( $m = 1, 2, \dots$ ), de la façon suivante :
  - a) Générer  $\xi \sim q(\xi|\theta^m)$
  - b) Poser

$$\varrho(\theta^{(m)}, \xi) = \min\left\{\frac{\pi(\xi)q(\theta^{(m)}|\xi)}{\pi(\theta^{(m)})q(\xi|\theta^{(m)})}, 1\right\}$$

c) prendre

$$\theta^{(m+1)} = \begin{cases} \xi & \text{avec probabilité } \varrho(\theta^{(m)}, \xi) \\ \theta^{(m)} & \text{sinon} \end{cases}$$

Une propriété stupéfiante de cet algorithme est d'autoriser toute chaîne de markov convergeant vers la loi d'intérêt. La loi de densité  $\pi(\theta|x)$  est souvent appelée *loi cible* ou *loi objet*, tandis que la loi de densité  $q(.|\theta)$  est dite *loi de proposition* tel que  $q$  doit vérifier certaines conditions telles qu'elle doit être *simulable*, son support doit couvrir le support de  $\pi$ , et doit être une bonne approximation de  $\pi$ . A consulter : [Robert et Cassella \(1999\)](#) ; [Gamerman, Dani, and Hedibert F. Lopes \(2006\)](#), [Robert et smith \(1993\)](#)

### 1.3.1.2 L'échantillonnage de Gibbs

La méthode *Metropolis-Hastings* est attrayante par son universalité, mais pratiquement peut empêcher la convergence si la probabilité d'atteindre des parties éloignées du support de la loi de  $\pi$  est trop petite.

L'algorithme de Gibbs est une méthode utilisée pour la première fois pour des problèmes d'échantillonnage dans le cadre de l'analyse bayésienne par les frères D.Geman et S.Geman dans les années 1980 puis ils l'ont introduit dans leur travail qui fut publié dans un article intitulé "Stochastic Relaxation, Gibbs Distributions, and the Bayesian Restoration of Images" en (1984) pour générer des observations à partir d'une distribution de Gibbs (Distribution de Boltzman). Il s'agit d'une forme particulière de la méthode Monte-Carlo par chaîne de Markov qui par son efficacité est largement utilisée dans de nombreux domaines de l'analyse bayésienne. Une bonne introduction à l'échantillonnage de Gibbs est

donnée par [Cassella \(1992\)](#).

Dans la méthode de Gibbs, après avoir choisi un point de départ, les  $d$  composantes du vecteur  $(\theta)$  sont générées les unes après les autres conditionnellement à toutes les autres composantes. Si  $\pi(\theta|x)$  est la densité des  $d$  composantes du vecteur  $\theta$ , conditionnellement aux données observées  $(x)$ , nous utilisons alors les densités conditionnelles  $\pi(\theta_1|\theta_2, \theta_3, \dots, \theta_d, x)$ ,  $\pi(\theta_2|\theta_1, \theta_3, \dots, \theta_d, x)$ , et ainsi de suite. À chaque  $k^e$  étape, la distribution conditionnelle utilise les valeurs générées les plus récentes parmi toutes les autres composantes. Par la théorie des chaînes de Markov, lorsque  $k \rightarrow \infty$  la densité des réalisations obtenues converge vers  $\pi(\theta|x)$ . Le schéma de simulation est décrit ci-dessous :

***Algorithme d'échantillonnage de Gibbs***

1. fixer  $k=0$
2. — Générer  $\theta_1^{(k+1)}$  à partir de  $\pi(\theta_1|\theta_2^{(k)}, \theta_3^{(k)}, \dots, \theta_d^{(k)}, x)$ ,  
     — Générer  $\theta_2^{(k+1)}$  à partir de  $\pi(\theta_2|\theta_1^{(k+1)}, \theta_3^{(k)}, \dots, \theta_d^{(k)}, x)$ ,  
     ...  
     — Générer  $\theta_{d-1}^{(k+1)}$  à partir de  $\pi(\theta_{d-1}|\theta_1^{(k+1)}, \theta_2^{(k+1)}, \dots, \theta_d^{(k)}, x)$ ,  
     — Générer  $\theta_d^{(k+1)}$  à partir de  $\pi(\theta_d|\theta_1^{(k+1)}, \theta_2^{(k+1)}, \dots, \theta_{d-1}^{(k+1)}, x)$ ,
3. Si la convergence est obtenue, alors
  - 2.a Retenir  $\theta = \theta^{(k+1)}$   
     Sinon
  - 2.b Fixer  $k = k + 1$ , et retourner à 1

Le schéma de simulation est d'autant plus efficace lorsque les distributions

conditionnelles  $\pi(\theta_i|\theta_1, \dots, \theta_{(i-1)}, \theta_{(i+1)}, \dots, \theta_d)$  sont simulables rapidement.

Un des points remarquables de l'algorithme de Gibbs est qu'il rend possible la simulation de la distribution jointe  $\pi(\theta) = \pi(\theta_1, \dots, \theta_d)$  à partir des distributions conditionnelles  $\pi_i(\theta_i|\theta_{j \neq i})$ .

La convergence de l'algorithme de Gibbs fait référence à la propriété selon laquelle, à mesure que le nombre d'itérations de l'algorithme augmente, les échantillons générés convergent vers une approximation de la distribution cible. Plus formellement, l'algorithme de Gibbs converge si, pour un nombre suffisamment grand d'itérations, les échantillons générés suivent la distribution cible avec une bonne précision. A consulter [Gilks, Walter R., Sylvia Richardson, and David Spiegelhalter \(1995\)](#), [Gelman, Andrew \(2013\)](#), [Robert et Casella \(1999\)](#)

## 1.4 Modèles autorégressifs (AR)

**Définition 1.8.** *Un modèle autorégressif d'ordre  $p$  ( $AR(p)$ ) est défini comme suit*

$$y_t = c + \sum_{i=1}^p \phi_i y_{t-i} + \epsilon_t$$

où

- $y_t$  est la variable dépendante à un moment donné  $t$
- $c$  est une constante (ou un intercept)
- $\phi_i$  sont les coefficients autorégressifs qui représentent l'influence des  $p$  valeurs passées de la variable dépendante sur sa valeur actuelle. Ces coefficients peuvent être positifs ou négatifs.
- $\epsilon_t$  est le terme d'erreur à un moment donné  $t$ .

**Exemple 1.5.** *On prend l'exemple d'un  $AR(2)$  prenant la forme suivante*

$$y_t = c + \phi_1 y_{t-1} + \phi_2 y_{t-2} + \epsilon_t$$

*Ce modèle AR(2) peut être utilisé pour prédire les valeurs futures de la série temporelle  $y_t$  en utilisant les valeurs passées  $y_{t-1}$  et  $y_{t-2}$ , ainsi que les coefficients autorégressifs  $\phi_1$  et  $\phi_2$  estimés à partir des données historiques de la série temporelle.*

# Chapitre 2

## Inférence Bayésienne d'un modèle de régression linéaire sous contamination

### 2.1 Introduction

L'estimation bayésienne des paramètres d'un modèle de régression linéaire sous contamination vise à estimer les paramètres du modèle de régression linéaire en tenant compte de la présence de données contaminées ou aberrantes. La contamination fait référence à la présence de valeurs atypiques ou d'erreurs dans les données qui peuvent affecter les estimations des paramètres du modèle.

### 2.2 Régression linéaire simple

La régression linéaire simple est un modèle de régression linéaire avec une seule variable explicative. Elle permet de résumer et d'étudier la relation entre deux variables continues (quantitatives). Dans ce type de modèle, une variable

dépendante est expliquée par une seule variable indépendante représentée par l'équation suivante :

$$Y_i = \beta_0 + \beta_1 X_i + \epsilon_i \quad i = 1, \dots, n$$

avec

- $Y_i$  est la variable dépendante (la variable à prédire)
- $X_i$  est la variable indépendante (la variable utilisée pour prédire  $Y$ )
- $\beta_0$  est l'intercept (la valeur de  $Y$  lorsque  $X$  est égale à zéro)
- $\beta_1$  est la pente (la variation de  $Y$  pour une variation d'une unité de  $X$ )
- $\epsilon_i$  est le terme d'erreur (la différence entre la valeur prédite de  $Y$  et la valeur réelle de  $Y$ )
- $n$  est le nombre d'observations.

L'objectif de la régression linéaire simple est de trouver les valeurs de  $\beta_0$  et  $\beta_1$  qui minimisent la somme des carrés des écarts entre les valeurs prédites et les valeurs réelles de  $Y$ . Cette méthode est souvent utilisée pour comprendre la relation entre deux variables et pour prédire les valeurs de la variable dépendante à partir des valeurs de la variable indépendante.

### 2.2.1 Hypothèses du modèle

Les principales hypothèses sur lesquelles repose le modèle sont les suivantes :

- 1)  $E(\epsilon) = 0$ , l'erreur est centrée.
- 2)  $E(\epsilon^2) = \sigma^2$ , la variance de l'erreur est constante.
- 3)  $cov(\epsilon_i, \epsilon_j) = 0$ , si  $\epsilon_i \neq \epsilon_j$ , les erreurs ne sont pas autocorrélées.
- 4)  $cov(X_i, \epsilon_j) = 0$ , l'erreur n'est pas corrélé avec la variable indépendante.
- 5) La variable exogène n'est pas aléatoire.

## 2.3 Estimation des paramètres d'un modèle de régression linéaire quand les erreurs sont dépendantes

Le modèle de régression avec erreurs autocorrélées est l'un des modèles les plus fortement analysés en économétrie. Bien que les idées fréquentistes soient dominantes, il existe une quantité assez importante de travaux bayésiens dans ce domaine Zellner and Tiao (1964), Zellner (1971), Hamilton (2020), Judge, G. G (1988) .

Dans ce chapitre nous allons nous intéresser au cas où les erreurs  $\mu_i$  sont dépendantes (autocorrélées) (Voir Ait mohammed (2006)).

### 2.3.1 Le modèle

On considère le modèle suivant

$$\begin{cases} Y_t = \beta X_t + \mu_t & t = 1 \dots n \\ \mu_t = \alpha \mu_{t-1} + \epsilon_t \end{cases} \quad (2.1)$$

où les  $\epsilon_t$  sont des variables aléatoires indépendantes identiquement distribuées (iid) de loi normale  $N(0, \sigma^2)$ , les paramètres du modèle sont supposés inconnus,  $\beta, \alpha \in \mathbb{R}$  et  $\sigma > 0$ . Nous supposons que le processus  $AR(1)$  est stationnaire.

En supposant que (2.1) est représentative de ce qui s'est passé pour  $t = 0, 1, 2, \dots, n_0$ , où  $n_0$  est inconnu, nous avons, par exemple, pour  $t = 0$  en supposant  $\epsilon_0 \sim \mathcal{N}(0, \sigma^2)$

$$y_0 = x_0 \beta + A + \mu_0$$

avec  $A = \alpha(y_{-1} - \theta x_{-1})$  une fonction de quantités non observables et non observées. Sous ces hypothèses,  $y_0 \sim \mathcal{N}(\beta x_0 + A, \sigma^2)$ .

### Fonction de vraisemblance

En remplaçant la formule de  $\mu_t$  dans  $y_t$  et  $\mu_{t-1}$  par sa valeur en fonction de  $y_{t-1}$ , on obtient

$$\begin{aligned} y_t &= \beta x_t + \alpha \mu_{t-1} + \epsilon_t \\ &= \beta x_t + \alpha(y_{t-1} - \beta x_{t-1}) + \epsilon_t \end{aligned}$$

La fonction de vraisemblance s'écrit comme suit

$$L(\theta|y) \propto \frac{1}{\sigma^n} \exp\left\{-\frac{1}{2\sigma^2} \sum_{t=1}^n (y_t - \beta x_t - \alpha(y_{t-1} - \beta x_{t-1}))^2\right\} \quad (2.2)$$

### 2.3.2 Lois a posteriori conditionnelles de $\beta$ et $\alpha$

La fonction de vraisemblance des paramètres  $\beta, \alpha, \sigma$  et  $A$  est donnée par :

$$L(\beta, \alpha, \sigma|y) \propto \frac{1}{\sigma^{-(n+1)}} \exp\left\{-\frac{1}{2\sigma^2} (y_0 - \beta x_0 - A)^2\right\} \exp\left\{-\frac{1}{2\sigma^2} \sum_{t=1}^n (y_t - \alpha y_{t-1} - \beta(x_t - \alpha x_{t-1}))^2\right\} \quad (2.3)$$

Selon la formule de Bayes, la loi a posteriori de  $(\beta, \alpha, \sigma, A)$  sachant les observations  $y, x$  est proportionnelle au produit de la vraisemblance et la loi a priori. D'où la loi a posteriori des paramètres est donnée par :

$$\pi(\beta, \alpha, \sigma, A|y, x) \propto L(\beta, \alpha, \sigma, A|y) \pi(\beta, \alpha, \sigma, A) \quad (2.4)$$

On suppose que les paramètres  $\beta, \alpha, \sigma, A$  sont indépendants et de loi a priori non informative :

$$\pi(\beta, \alpha, \sigma, A) \propto \frac{1}{\sigma} \quad (2.5)$$

En remplaçant (2.3) et (2.5) dans (2.4), on obtient :

$$\pi(\beta, \alpha, \sigma, A|y, x) \propto \frac{1}{\sigma^{n+2}} \exp\left\{ \frac{-1}{2\sigma^2} \left[ (y_0 - \beta x_0 - A)^2 + \sum_{t=1}^n (y_t - \alpha y_{t-1} - \beta(x_t - \alpha x_{t-1}))^2 \right] \right\} \quad (2.6)$$

En intégrant par rapport à  $A$  et en tenant compte des propriétés de la loi normale, on trouve :

$$\begin{aligned} \pi(\beta, \alpha, \sigma|y, x) &\propto \int_{-\infty}^{+\infty} \pi(\beta, \alpha, \sigma, A|y, x) dA \\ &\propto \frac{1}{\sigma^{n+2}} \exp\left\{ -\frac{1}{2\sigma^2} \left[ (y_0 - \beta x_0 - A)^2 + \sum_{t=1}^n (y_t - \alpha y_{t-1} - \beta(x_t - \alpha x_{t-1}))^2 \right] \right\} \end{aligned}$$

d'où

$$\pi(\beta, \alpha, \sigma|y, x) \propto \frac{1}{\sigma^{n+1}} \exp\left\{ \frac{-1}{2\sigma^2} \sum_{t=1}^n (y_t - \alpha y_{t-1} - \beta(x_t - \alpha x_{t-1}))^2 \right\}$$

car

$$\int_{-\infty}^{+\infty} \frac{1}{\sqrt{2\pi}\sigma} \exp\left\{ -\frac{1}{2\sigma^2} (y_0 - \beta x_0 - A)^2 \right\} dA = 1$$

L'intégration par rapport à  $\sigma$ , en prenant en considération les propriétés de la loi gamma donne  $\pi(\beta, \alpha|y, x)$ .

$$\begin{aligned} \pi(\beta, \alpha|y, x) &\propto \int_0^{+\infty} \pi(\beta, \alpha, \sigma|y) d\sigma \\ &\propto \int_0^{+\infty} \frac{1}{\sigma^{n+1}} \exp\left\{ \frac{-1}{2\sigma^2} \sum_{t=1}^n (y_t - \alpha y_{t-1} - \beta(x_t - \alpha x_{t-1}))^2 \right\} d\sigma \\ &\propto \left\{ \sum_{t=1}^n (y_t - \alpha y_{t-1} - \beta(x_t - \alpha x_{t-1}))^2 \right\}^{-\frac{n}{2}} \\ &= \left\{ \sum_{t=1}^n (y_t - \beta x_t - \alpha(y_{t-1} - \beta x_{t-1}))^2 \right\}^{-\frac{n}{2}} \quad (2.7) \end{aligned}$$

Afin d'utiliser le Gibbs sampler, on aura besoin de  $\pi(\beta|\alpha, y, x)$  et  $\pi(\alpha|\beta, y, x)$ .

### 2.3.2.1 Loi a posteriori conditionnelle de $\beta$

$$\pi(\beta|\alpha, y, x) \propto \left\{ \sum_{t=1}^n (y_t - \alpha y_{t-1})^2 - 2\beta \sum_{t=1}^n (y_t - \alpha y_{t-1})(x_t - \alpha x_{t-1}) + \beta^2 \sum_{t=1}^n (x_t - \alpha x_{t-1})^2 \right\} \quad (2.8)$$

On pose

$$\begin{aligned} a &= \sum_{t=1}^n (y_t - \alpha y_{t-1})^2, \\ b &= \sum_{t=1}^n (y_t - \alpha y_{t-1})(x_t - \alpha x_{t-1}) \\ c &= \sum_{t=1}^n (x_t - \alpha x_{t-1})^2 \end{aligned} \quad (2.9)$$

On obtient

$$\begin{aligned} \pi(\beta|\alpha, y, x) &\propto [a - 2\beta b + c\beta^2]^{-\frac{n}{2}} \\ &\propto \left( \frac{a}{c} - 2\frac{b\beta}{c} + \beta^2 \right)^{-\frac{n}{2}} \\ &\propto \left[ 1 + \frac{(\beta - \frac{b}{c})^2}{(n-1)\frac{(\frac{a}{c} - \frac{b^2}{c^2})}{(n-1)}} \right]^{-\frac{n}{2}} \end{aligned}$$

qui est la loi a posteriori conditionnelle de  $\beta$ .

$$\pi(\beta|\alpha, y, x) \propto \left[ 1 + \frac{(\beta - \hat{\beta}(\alpha))^2}{(n-1)K(\alpha)} \right]^{-\frac{n}{2}} \quad (2.10)$$

Avec

$$\hat{\beta}(\alpha) = \frac{b}{c} = \frac{\sum_{t=1}^n (y_t - \alpha y_{t-1})(x_t - \alpha x_{t-1})}{\sum_{t=1}^n (x_t - \alpha x_{t-1})^2}$$

Et

$$\begin{aligned} K(\alpha) &= \frac{1}{n-1} \left( \frac{a}{c} - \frac{b^2}{c^2} \right) \\ &= \frac{1}{n-1} \left[ \frac{\sum_{t=1}^n (y_t - \alpha y_{t-1})^2}{\sum_{t=1}^n (x_t - \alpha x_{t-1})^2} - \frac{(\sum_{t=1}^n (y_t - \alpha y_{t-1})(x_t - \alpha x_{t-1}))^2}{\sum_{t=1}^n (x_t - \alpha x_{t-1})^2} \right] \end{aligned}$$

On obtient le noyau d'une loi de student de moyenne  $\hat{\beta}(\alpha)$ , d'une variance  $K(\alpha)$  et de  $(n-1)$  degrés de liberté.

$$\beta | \alpha, y, x \sim t(\hat{\beta}(\alpha), K(\alpha), n-1)$$

### 2.3.2.2 Loi a posteriori conditionnelle de $\alpha$

$$\pi(\alpha | \beta, y, x) \propto \left\{ \sum_{t=1}^n (y_t - \beta x_t)^2 - 2\alpha \sum_{t=1}^n (y_t - \beta x_t)(y_{t-1} - \beta x_{t-1}) + \alpha^2 \sum_{t=1}^n (y_{t-1} - \beta x_{t-1})^2 \right\}$$

On pose

$$\begin{aligned} a' &= \sum_{t=1}^n (y_t - \beta x_t)^2, \\ b' &= \sum_{t=1}^n (y_t - \beta x_t)(y_{t-1} - \beta x_{t-1}) \\ c' &= \sum_{t=1}^n (y_{t-1} - \beta x_{t-1})^2 \end{aligned}$$

On obtient

$$\begin{aligned}
 \pi(\alpha|\beta, y, x) &\propto [a' - 2\alpha b' + c'\alpha^2]^{-\frac{n}{2}} \\
 &\propto \left(\frac{a'}{c'} - 2\frac{b'\alpha}{c'} + \alpha^2\right)^{-\frac{n}{2}} \\
 &\propto \left[1 + \frac{(\alpha - \frac{b'}{c'})^2}{(n-1)\frac{(\frac{a'}{c'} - \frac{b'^2}{c'^2})}{(n-1)}}\right]^{-\frac{n}{2}}
 \end{aligned}$$

Par conséquent

$$\pi(\alpha|\beta, y, x) \propto \left[1 + \frac{(\alpha - \hat{\alpha}(\beta))^2}{(n-1)K(\beta)}\right]^{-\frac{n}{2}} \quad (2.11)$$

qui est le noyau d'une loi de student de moyenne  $\hat{\alpha}(\beta)$ , d'une variance  $K(\beta)$  et de  $(n-1)$  degrés de liberté.

$$\alpha|\beta, y, x \sim t(\hat{\alpha}(\beta), K(\beta), n-1)$$

Avec

$$\hat{\alpha}(\beta) = \frac{b'}{c'} = \frac{\sum_{t=1}^n (y_t - \beta y_{t-1})(x_t - \beta x_{t-1})}{\sum_{t=1}^n (x_t - \beta x_{t-1})^2}$$

et

$$\begin{aligned}
 K(\beta) &= \frac{1}{n-1} \left(\frac{a'}{c'} - \frac{b'^2}{c'^2}\right) \\
 &= \frac{1}{n-1} \left[ \frac{\sum_{t=1}^n (y_t - \beta y_{t-1})^2}{\sum_{t=1}^n (y_{t-1} - \beta x_{t-1})^2} - \left(\frac{\sum_{t=1}^n (y_t - \beta y_{t-1})(x_t - \alpha x_{t-1})}{\sum_{t=1}^n (y_{t-1} - \beta x_{t-1})^2}\right)^2 \right]
 \end{aligned}$$

## 2.4 Estimation des paramètres d'un modèle de régression linéaire quand les erreurs sont dépendantes ( $\epsilon_t \sim AR(1)$ ) sous contamination AO

Nous allons déterminer les distributions *a posteriori* conditionnelles des paramètres du modèle de régression quand les erreurs sont autocorrélées sous contamination AO. A consulter pour plus de détails à ce sujet : [Box, George EP \(1968\)](#); [Chaloner, Kathryn et Rollin Brant \(1988\)](#); [Jack, Gambino et Guttman Irwin \(1984\)](#), [Jack Gambino et Guttman Irwin \(1984\)](#), [Chen et Cathy WS \(1997\)](#); [Galeano, Pedro, and Daniel Peña \(2013\)](#)

### 2.4.1 Présentation du modèle AO

Dans ce cas une seule observation est affectée, au lieu d'observer les  $y_t$ , nous observons  $y_t^*$

$$\begin{aligned}
 y_t^* &= \begin{cases} y_t & si & t \neq k \\ y_t + \Delta & si & t = k \end{cases} \\
 &= \begin{cases} \beta x_t + \mu_t & t \neq k \\ \mu_t = \alpha \mu_{t-1} + \epsilon_t & \\ \beta x_k + \mu_k + \Delta & t = k \end{cases} \quad (2.12)
 \end{aligned}$$

avec  $\Delta$  est l'amplitude inconnue de la contamination qui se produit à l'instant  $t = k$  avec  $1 < k < n$ .

### 2.4.2 Lois a posteriori conditionnelles de $\beta, \alpha, \Delta$

$$\begin{aligned}
 y_{k+1} &= \beta x_{k+1} + \mu_{k+1} \\
 &= \beta x_{k+1} + \alpha \mu_k + \epsilon_{k+1} \\
 &= \beta x_{k+1} + \alpha(y_k - \beta x_k - \Delta) + \epsilon_{k+1} \\
 &= \beta x_{k+1} + \alpha y_k - \alpha \beta x_k - \alpha \Delta + \epsilon_{k+1} \\
 y_{k+1} - \alpha y_k &= \beta(x_{k+1} - \alpha x_k) - \alpha \Delta + \epsilon_{k+1}
 \end{aligned}$$

Le modèle peut être réécrit comme suit :

$$\begin{cases}
 y_t - \alpha y_{t-1} = \beta(x_t - \alpha x_{t-1}) + \epsilon_t, & t \neq k, k+1 \\
 y_k - \alpha y_{k-1} = \beta(x_k - \alpha x_{k-1}) + \epsilon_k + \Delta \\
 y_{k+1} - \alpha y_k = \beta(x_{k+1} - \alpha x_k) + \epsilon_{k+1} - \alpha \Delta
 \end{cases}$$

La loi a posteriori de  $(\beta, \alpha, \Delta, \sigma)$  est donnée comme suit

$$\begin{aligned}
 \pi(\beta, \alpha, \sigma, \Delta | y) &= \pi(\beta, \alpha, \sigma, \Delta) \cdot \pi(y_1, \dots, y_n | \beta, \alpha, \sigma, \Delta) \\
 &\propto (\sigma)^{-(n+1)} \exp\left\{ \frac{-1}{2\sigma^2} \left[ \sum_{\substack{t=1 \\ t \neq k, k+1}}^n (y_t - \alpha y_{t-1} - \beta(x_t - \alpha x_{t-1}))^2 \right. \right. \\
 &\quad \left. \left. + (y_k - \alpha y_{k-1} - \beta(x_k - \alpha x_{k-1}) - \Delta)^2 \right. \right. \\
 &\quad \left. \left. + (y_{k+1} - \alpha y_k - \beta(x_{k+1} - \alpha x_k) + \alpha \Delta)^2 \right] \right\} \quad (2.13)
 \end{aligned}$$

On introduit la notation

$$\begin{aligned}
 A(\beta, \alpha, \Delta) &= \sum_{\substack{t=1 \\ t \neq k, k+1}}^n (y_t - \alpha y_{t-1} - \beta(x_t - \alpha x_{t-1}))^2 \\
 &\quad + (y_k - \alpha y_{k-1} - \beta(x_k - \alpha x_{k-1}) - \Delta)^2 \\
 &\quad + (y_{k+1} - \alpha y_k - \beta(x_{k+1} - \alpha x_k) + \alpha \Delta)^2 \quad (2.14)
 \end{aligned}$$

Alors

$$\pi(\beta, \alpha, \sigma, \Delta|y) \propto (\sigma)^{-(n+1)} \exp\left\{\frac{-1}{2\sigma^2} A(\beta, \alpha, \Delta)\right\} \quad (2.15)$$

En effectuant le changement de variables  $z = \frac{1}{2\sigma^2} A(\beta, \alpha, \Delta)$  et en intégrant par rapport à  $\sigma$  dans (2.15), on obtient la loi a posteriori conditionnelle suivante

$$\begin{aligned} \pi(\beta, \alpha, \Delta|y, x) &\propto \int_0^{+\infty} A(\beta, \alpha, \Delta)^{-\frac{n}{2}} z^{\frac{n}{2}-1} \exp(-z) dz \\ &\propto \{A(\beta, \alpha, \Delta)\}^{-\frac{n}{2}} \\ &\propto \left\{ \sum_{t \neq k, k+1} (y_t - \alpha y_{t-1} - \beta(x_t - \alpha x_{t-1}))^2 \right. \\ &\quad \left. + (y_k - \alpha y_{k-1} - \beta(x_k - \alpha x_{k-1}) - \Delta)^2 \right. \\ &\quad \left. + (y_{k+1} - \alpha y_k - \beta(x_{k+1} - \alpha x_k) + \alpha \Delta)^2 \right\}^{-\frac{n}{2}} \end{aligned}$$

#### 2.4.2.1 Loi a posteriori de $\beta$ sachant $\alpha$ et $\Delta$

la loi a posteriori de  $\beta$  pour les valeurs fixées de  $\alpha$  et  $\Delta$

$$\begin{aligned} \pi(\beta|\alpha, \Delta, y, x) &\propto \left\{ \sum_{t=1}^n (y_t - \alpha y_{t-1})^2 - 2\Delta((y_k - \alpha y_{k-1}) - 2\alpha(y_{k+1} - \alpha y_k)) + \Delta^2(1 + \alpha^2) \right. \\ &\quad \left. - 2\beta \left[ \sum_{t=1}^n (y_t - \alpha y_{t-1})(x_t - \alpha x_{t-1}) - \Delta(x_k - \alpha x_{k-1}) + \alpha \Delta(x_{k+1} - \alpha x_k) \right] \right. \\ &\quad \left. + \beta^2 \sum_{t=1}^n (x_t - \alpha x_{t-1})^2 \right\}^{-\frac{n}{2}} \end{aligned}$$

On pose

$$\begin{aligned} a_1(\alpha, \Delta) &= \sum_{t=1}^n (y_t - \alpha y_{t-1})^2 - 2\Delta(y_k - \alpha y_{k-1}) + 2\alpha \Delta(y_{k+1} - \alpha y_k) + \Delta^2(1 + \alpha^2) \\ b_1(\alpha, \Delta) &= \sum_{t=1}^n (y_t - \alpha y_{t-1})(x_t - \alpha x_{t-1}) - \Delta(x_k - \alpha x_{k-1}) + \alpha \Delta(x_{k+1} - \alpha x_k) \\ c_1(\alpha, \Delta) &= \sum_{t=1}^n (x_t - \alpha x_{t-1})^2 \end{aligned}$$

Ce qui nous permet d'écrire la densité  $\pi(\beta|\alpha, \Delta, y, x)$  sous la forme

$$\begin{aligned}\pi(\beta|\alpha, \Delta, y, x) &\propto \{a_1(\alpha, \Delta) - 2\beta b_1(\alpha, \Delta) + \beta^2 c_1(\alpha, \Delta)\}^{-\frac{n}{2}} \\ &= \left\{1 + \frac{(\beta - \hat{\beta}(\alpha, \Delta))^2}{S_1^2(\alpha, \Delta)(n-1)}\right\}^{-\frac{n}{2}}\end{aligned}$$

qui est une loi de student de moyenne  $\hat{\beta}(\alpha, \Delta)$  de variance  $S^2(\alpha, \Delta)$  et de  $(n-1)$  degrés de liberté avec

$$\begin{aligned}\hat{\beta}(\alpha, \Delta) &= \frac{b_1(\alpha, \Delta)}{c_1(\alpha, \Delta)} \\ &= \frac{\sum_{t=1}^n (y_t - \alpha y_{t-1})(x_t - \alpha x_{t-1}) - \Delta(x_k - \alpha x_{k-1}) + \alpha \Delta(x_{k+1} - \alpha x_k)}{\sum_{t=1}^n (x_t - \alpha x_{t-1})^2}\end{aligned}$$

$$\begin{aligned}S_1^2(\alpha, \Delta) &= \frac{1}{n-1} \left( \frac{a_1(\alpha, \Delta)}{c_1(\alpha, \Delta)} - \frac{b_1(\alpha, \Delta)^2}{c_1(\alpha, \Delta)^2} \right) \\ &= \frac{1}{n-1} \left( \frac{\sum_{t=1}^n (y_t - \alpha y_{t-1})^2 - 2\Delta(y_k - \alpha y_{k-1}) + 2\alpha \Delta(y_{k+1} - \alpha y_k) + \Delta^2(1 + \alpha^2)}{\sum_{t=1}^n (x_t - \alpha x_{t-1})^2} \right. \\ &\quad \left. - \frac{(\sum_{t=1}^n (y_t - \alpha y_{t-1})(x_t - \alpha x_{t-1}) - \Delta(x_k - \alpha x_{k-1}) + \alpha \Delta(x_{k+1} - \alpha x_k))^2}{(\sum_{t=1}^n (x_t - \alpha x_{t-1})^2)^2} \right)\end{aligned}$$

### 2.4.2.2 Loi a posteriori de $\alpha$ sachant $\Delta$ et $\beta$

La loi a posteriori de  $\alpha$  pour les valeurs fixées de  $\Delta$  et  $\beta$

$$\begin{aligned} \pi(\alpha|\beta, \Delta, y, x) \propto & \left\{ \sum_{t=1}^n (y_t - \beta x_t)^2 - 2\Delta(y_k - \beta x_k) + \Delta^2 \right. \\ & - 2\alpha \left[ \sum_{t=1}^n (y_t - \beta x_t)(y_{t-1} - \beta x_{t-1}) - \Delta[(y_{k-1} + \beta x_{k-1}) + (y_{k+1} - \beta x_{k+1})] \right] \\ & \left. + \alpha^2 \left[ \sum_{t=1}^n (y_{t-1} - \beta x_{t-1})^2 - 2\Delta(y_k - \beta x_k) + \Delta^2 \right] \right\}^{-\frac{n}{2}} \end{aligned}$$

On pose

$$\begin{aligned} a_2(\Delta, \beta) &= \sum_{t=1}^n (y_t - \beta x_t)^2 - 2\Delta(y_k - \beta x_k) + \Delta^2 \\ b_2(\Delta, \beta) &= \sum_{t=1}^n (y_t - \beta x_t)(y_{t-1} - \beta x_{t-1}) - \Delta[(y_{k-1} + \beta x_{k-1}) + (y_{k+1} - \beta x_{k+1})] \\ c_2(\Delta, \beta) &= \sum_{t=1}^n (y_{t-1} - \beta x_{t-1})^2 - 2\Delta(y_k - \beta x_k) + \Delta^2 \end{aligned}$$

Ce qui nous permet d'écrire la densité  $\pi(\alpha|\beta, \Delta, y, x)$  sous la forme

$$\begin{aligned} \pi(\alpha|\beta, \Delta, y, x) &= \{a_2(\Delta, \beta) - 2\alpha b_2(\Delta, \beta) + \alpha^2 c_2(\Delta, \beta)\}^{-\frac{n}{2}} \\ &= \left\{ 1 + \frac{(\alpha - \hat{\alpha}(\Delta, \beta))^2}{S_2^2(\Delta, \beta)(n-1)} \right\}^{-\frac{n}{2}} \end{aligned}$$

qui est une loi de student de moyenne  $\hat{\alpha}(\Delta, \beta)$  de variance  $S^2(\Delta, \beta)$  et de  $(n-1)$  degrés de liberté avec

$$\begin{aligned} \hat{\alpha}(\Delta, \beta) &= \frac{b_2(\Delta, \beta)}{c_2(\Delta, \beta)} \\ &= \frac{\sum_{t=1}^n (y_t - \beta x_t)(y_{t-1} + \beta x_{t-1}) - \Delta[(y_{k-1} + \beta x_{k-1}) + (y_{k+1} - \beta x_{k+1})]}{\sum_{t=1}^n (y_{t-1} - \beta x_{t-1})^2 - 2\Delta(y_k - \beta x_k) + \Delta^2} \end{aligned}$$

$$\begin{aligned}
 S_2^2(\Delta, \beta) &= \frac{1}{n-1} \left( \frac{a_2(\Delta, \beta)}{c_2(\Delta, \beta)} - \frac{b_2(\Delta, \beta)^2}{c_2(\Delta, \beta)^2} \right) \\
 &= \frac{1}{n-1} \left( \frac{\sum_{t=1}^n (y_t - \beta x_t)^2 - 2\Delta(y_k - \beta x_k) + \Delta^2}{\sum_{t=1}^n (y_{t-1} - \beta x_{t-1})^2 - 2\Delta(y_k - \beta x_k) + \Delta^2} \right. \\
 &\quad \left. - \frac{(\sum_{t=1}^n (y_t - \beta x_t)(y_{t-1} + \beta x_{t-1}) - \Delta[(y_{k-1} + \beta x_{k-1}) + (y_{k+1} - \beta x_{k+1})])^2}{(\sum_{t=1}^n (y_{t-1} - \beta x_{t-1})^2 - 2\Delta(y_k - \beta x_k) + \Delta^2)^2} \right)
 \end{aligned}$$

### 2.4.2.3 Loi a posteriori de $\Delta$ sachant $\alpha$ et $\beta$

La loi a posteriori de  $\Delta$  pour les valeurs fixées de  $\alpha$  et  $\beta$

$$\begin{aligned}
 \pi(\Delta|\alpha, \beta, y, x) &= \left\{ \sum_{\substack{t=1 \\ t \neq k, k+1}}^n (y_t - \alpha y_{t-1} - \beta(x_t - \alpha x_{t-1}))^2 \right. \\
 &\quad + (y_k - \alpha y_{k-1} - \beta(x_k - \alpha x_{k-1}))^2 + \Delta^2 \\
 &\quad - 2\Delta(y_k - \alpha y_{k-1} - \beta(x_k - \alpha x_{k-1})) \\
 &\quad + (y_{k+1} - \alpha y_k - \beta(x_{k+1} - \alpha x_k))^2 + \alpha^2 \Delta^2 \\
 &\quad \left. - 2\alpha \Delta(y_{k+1} - \alpha y_k - \beta(x_{k+1} - \alpha x_k)) \right\}^{-\frac{n}{2}}
 \end{aligned}$$

De là on trouve

$$\begin{aligned}
 \pi(\Delta|\alpha, \beta, y, x) &\propto \left\{ \sum_{\substack{t=1 \\ t \neq k, k+1}}^n (y_t - \alpha y_{t-1} - \beta(x_t - \alpha x_{t-1}))^2 \right. \\
 &\quad + (y_k - \alpha y_{k-1} - \beta(x_k - \alpha x_{k-1}))^2 + (y_{k+1} - \alpha y_k - \beta(x_{k+1} - \alpha x_k))^2 \\
 &\quad - 2\Delta[(y_k - \alpha y_{k-1} - \beta(x_k - \alpha x_{k-1})) + \alpha(y_{k+1} - \alpha y_k - \beta(x_{k+1} - \alpha x_k))] \\
 &\quad \left. + \Delta^2(1 + \alpha^2) \right\}^{-\frac{n}{2}}
 \end{aligned}$$

On pose

$$\begin{aligned}
 a_3(\alpha, \beta) &= \sum_{\substack{t=1 \\ t \neq k, k+1}}^n (y_t - \alpha y_{t-1} - \beta(x_t - \alpha x_{t-1}))^2 \\
 &\quad + (y_k - \alpha y_{k-1} - \beta(x_k - \alpha x_{k-1}))^2 + (y_{k+1} - \alpha y_k - \beta(x_{k+1} - \alpha x_k))^2 \\
 b_3(\alpha, \beta) &= (y_k - \alpha y_{k-1} - \beta(x_k - \alpha x_{k-1})) + \alpha(y_{k+1} - \alpha y_k - \beta(x_{k+1} - \alpha x_k)) \\
 c_3(\alpha) &= (1 + \alpha^2)
 \end{aligned}$$

Ce qui nous permet d'écrire la densité  $\pi(\Delta|\alpha, \beta, y, x)$  sous la forme

$$\begin{aligned}
 \pi(\Delta|\alpha, \beta, y) &= \{a_3(\alpha, \beta) - 2b_3(\alpha, \beta)\Delta + c_3(\alpha)\Delta^2\}^{-\frac{n}{2}} \\
 &= \left\{1 + \frac{(\Delta - \hat{\Delta}(\alpha, \beta))^2}{S_3^2(\alpha, \beta)(n-1)}\right\}^{-\frac{n}{2}}
 \end{aligned}$$

qui est une loi de student de moyenne  $\hat{\Delta}(\alpha, \beta)$  de variance  $S_3^2(\alpha, \beta)$  et de  $(n-1)$  degrés de liberté.

$$\begin{aligned}
 \hat{\Delta}(\alpha, \beta) &= \frac{b_3(\alpha, \beta)}{c_3(\alpha)} \\
 &= \frac{(y_k - \alpha y_{k-1} - \beta(x_k - \alpha x_{k-1})) + \alpha(y_{k+1} - \alpha y_k - \beta(x_{k+1} - \alpha x_k))}{(1 + \alpha^2)}
 \end{aligned}$$

$$\begin{aligned}
 S_3^2(\alpha, \beta) &= \frac{1}{n-1} \left( \frac{a_3(\alpha, \beta)}{c_3(\alpha, \beta)} - \frac{b_3(\alpha, \beta)^2}{c_3(\alpha, \beta)^2} \right) \\
 &= \frac{1}{n-1} \left( \frac{\sum_{\substack{t=1 \\ t \neq k, k+1}}^n (y_t - \alpha y_{t-1} - \beta(x_t - \alpha x_{t-1}))^2}{(1 + \alpha^2)} \right. \\
 &\quad \left. + \frac{(y_k - \alpha y_{k-1} - \beta(x_k - \alpha x_{k-1}))^2 + (y_{k+1} - \alpha y_k - \beta(x_{k+1} - \alpha x_k))^2}{(1 + \alpha^2)} \right. \\
 &\quad \left. - \frac{((y_k - \alpha y_{k-1} - \beta(x_k - \alpha x_{k-1})) + \alpha(y_{k+1} - \alpha y_k - \beta(x_{k+1} - \alpha x_k)))^2}{(1 + \alpha^2)^2} \right)
 \end{aligned}$$

## 2.5 Densité prédictive d'une observation future sous contamination AO

On définit la densité prédictive (*posterior predictive*) comme la densité d'une observation nouvelle de l'échantillon. On la calcule comme la densité de  $y_{n+1}$  sachant le vecteur  $(y_1, y_2, \dots, y_n)$  marginalement par rapport à  $\beta, \alpha, \Delta$  et  $y$ .

Dans notre cas, nous considérons  $y_{n+1}$  comme un paramètre inconnu. On cherchera alors à prédire  $y_{n+1}$  sachant  $(y_1, y_2, \dots, y_n)$  en calculant la loi a posteriori conditionnelle complète  $\pi(y_{n+1}|y, x, \sigma, \beta, \alpha, \Delta)$

On a

$$\pi(\beta, \sigma, \alpha, \Delta) \propto \frac{1}{\sigma} \quad (2.16)$$

et

$$\pi(y_{n+1}, \alpha, \beta, \Delta, \sigma) = \frac{1}{\sigma} \times \frac{1}{\sigma} \times \exp\left\{-\frac{1}{2\sigma^2}(y_{n+1} - \beta x_{n+1} - \alpha(y_n - \beta x_n))^2\right\}$$

avec  $y_{n+1} \sim \mathcal{N}(\beta x_{n+1} + \alpha(y_n - \beta x_n), \sigma^2)$

Ainsi que la distribution conjointe a posteriori des paramètres

$$\begin{aligned} \pi(y_{n+1}, \beta, \alpha, \Delta, \sigma|y, x_{n+1}, x) &\propto \frac{1}{\sigma^2} \exp\left\{-\frac{1}{2\sigma^2}(y_{n+1} - \beta x_{n+1} - \alpha(y_n - \beta x_n))^2\right\} \\ &\times \frac{1}{\sigma^n} \exp\left\{-\frac{1}{2\sigma^2} \sum_{\substack{t=1 \\ t \neq k, k+1}}^n (y_t - \beta x_t - \alpha(y_{t-1} - \beta x_{t-1}))^2\right. \\ &\quad \left.+ (y_k - \beta x_k - \Delta - \alpha(x_k - \beta x_{k-1}))^2\right. \\ &\quad \left.+ (y_{k+1} - \beta x_{k+1} + \alpha\Delta - \alpha(x_{k+1} - \beta x_k))^2\right\} \end{aligned}$$

d'où la loi a posteriori conditionnelle suivante

$$\begin{aligned} \pi(y_{n+1}|\beta, \alpha, \sigma, \Delta, y, x_{n+1}, x) &\propto \frac{1}{\sigma^{n+2}} \exp\left\{ -\frac{1}{2\sigma^2} \sum_{\substack{t=1 \\ t \neq k, k+1}}^n (y_t - \beta x_t - \alpha(y_{t-1} - \beta x_{t-1}))^2 \right. \\ &\quad + (y_k - \beta x_k - \Delta - \alpha(x_k - \beta x_{k-1}))^2 \\ &\quad \left. + (y_{k+1} - \beta x_{k+1} + \alpha\Delta - \alpha(x_{k+1} - \beta x_k))^2 \right\} \\ &\quad \times \exp\left\{ -\frac{1}{2\sigma^2} (y_{n+1} - \beta x_{n+1} - \alpha(y_n - \beta x_n))^2 \right\} \end{aligned}$$

### 2.5.1 Lois a posteriori conditionnelles de $y_{n+1}, \beta, \alpha, \Delta$

En intégrant par rapport à  $\sigma$  on trouve

$$\begin{aligned} \pi(y_{n+1}|\beta, \alpha, \Delta, y, x_{n+1}, x) &\propto \left[ \sum_{\substack{t=1 \\ t \neq k, k+1}}^n (y_t - \beta x_t - \alpha(y_{t-1} - \beta x_{t-1}))^2 \right. \\ &\quad + (y_k - \beta x_k - \Delta - \alpha(x_k - \beta x_{k-1}))^2 \\ &\quad + (y_{k+1} - \beta x_{k+1} + \alpha\Delta - \alpha(x_{k+1} - \beta x_k))^2 \\ &\quad \left. + (y_{n+1} - \beta x_{n+1} - \alpha(y_n - \beta x_n))^2 \right]^{-\frac{n+1}{2}} \end{aligned} \quad (2.17)$$

#### 2.5.1.1 Loi a posteriori conditionnelle de $y_{n+1}$

De (2.17) on obtient

$$\begin{aligned} \pi(y_{n+1}|\beta, \alpha, \Delta, y, x_{n+1}, x) &\propto \left[ \sum_{t=1}^n (y_t - \beta x_t - \alpha(y_{t-1} - \beta x_{t-1}))^2 \right. \\ &\quad + (y_k - \beta x_k - \Delta - \alpha(x_k - \beta x_{k-1}))^2 \\ &\quad + (y_{k+1} - \beta x_{k+1} + \alpha\Delta - \alpha(x_{k+1} - \beta x_k))^2 \\ &\quad + (\beta x_{n+1} + \alpha(y_n - \beta x_n))^2 - 2y_{n+1}(\beta x_{n+1} + \alpha(y_n - \beta x_n)) \\ &\quad \left. + y_{n+1}^2 \right]^{-\frac{n+1}{2}} \end{aligned}$$

En posant

$$\begin{aligned}
 a_1 &= \sum_{\substack{t=1 \\ t \neq k, k+1}}^n (y_t - \beta x_t - \alpha(y_{t-1} - \beta x_{t-1}))^2 + (y_k - \beta x_k - \Delta - \alpha(x_k - \beta x_{k-1}))^2 \\
 &\quad + (y_{k+1} - \beta x_{k+1} + \alpha\Delta - \alpha(x_{k+1} - \beta x_k))^2 + (\beta x_{n+1} + \alpha(y_n - \beta x_n))^2 \\
 a_2 &= \beta x_{n+1} + \alpha(y_n - \beta x_n) \\
 a_3 &= 1
 \end{aligned}$$

On obtient alors la loi a posteriori conditionnelle de  $y_{n+1}$  suivante

$$\begin{aligned}
 \pi(y_{n+1} | \alpha, \beta, \Delta, x, x_{n+1}) &\propto \{a_1 - 2a_2 y_{n+1} + a_3 y_{n+1}^2\}^{-\frac{n+1}{2}} \\
 &\propto \left[ 1 + \frac{(y_{n+1} - \frac{a_2}{a_3})^2}{n \frac{(\frac{a_1}{a_3} - \frac{a_2^2}{a_3^2})}{n}} \right]
 \end{aligned}$$

On obtient le noyau d'une loi de student telle que

$$y_{n+1} | \beta, \alpha, \Delta, y, X, x_{n+1} \sim t_n(\hat{y}_{n+1}, S_1, n)$$

avec

$$\begin{aligned}
 \hat{y}_{n+1} &= \frac{a_2}{a_3} \\
 &= \beta x_{n+1} + \alpha(y_n - \beta x_n)
 \end{aligned}$$

Et

$$\begin{aligned}
 S_1^2 &= \frac{1}{n} \left( \frac{a_1}{a_3} - \frac{a_2^2}{a_3^2} \right) \\
 &= \frac{1}{n} \left( \sum_{\substack{t=1 \\ t \neq k, k+1}}^n (y_t - \beta x_t - \alpha(y_{t-1} - \beta x_{t-1}))^2 + (y_k - \beta x_k - \Delta - \alpha(x_k - \beta x_{k-1}))^2 \right. \\
 &\quad \left. + (y_{k+1} - \beta x_{k+1} + \alpha\Delta - \alpha(x_{k+1} - \beta x_k))^2 \right)
 \end{aligned}$$

2.5.1.2 Loi a posteriori conditionnelle de  $\alpha$

$$\begin{aligned} \pi(\alpha|\beta, \Delta, y, x, x_{n+1}) &= \left[ \sum_{t=1}^n (y_t - \beta x_t)^2 + \Delta^2 - 2\Delta((y_k - \beta x_k) + (y_{k+1} - \beta x_{k+1})) \right. \\ &\quad - 2\alpha \left[ \sum_{t=1}^{n+1} (y_{t-1} - \beta x_{t-1})(y_t - \beta x_t) - \Delta[(x_k - \beta x_{k-1}) - (x_{k+1} - \beta x_k)] \right] \\ &\quad \left. + \alpha^2 \left[ \sum_{t=1}^{n+1} (y_{t-1} - \beta x_{t-1})^2 + \Delta^2 \right] + (y_{n+1} - \beta x_{n+1} - \alpha(y_n - \beta x_n))^2 \right]^{-\frac{n+1}{2}} \end{aligned}$$

On pose

$$\begin{aligned} b_1 &= \sum_{t=1}^n (y_t - \beta x_t)^2 + \Delta^2 - 2\Delta[(y_k - \beta x_k) + (y_{k+1} - \beta x_{k+1})] \\ &\quad + (y_{n+1} - \beta x_{n+1} - \alpha(y_n - \beta x_n))^2 \\ b_2 &= \sum_{t=1}^n (y_{t-1} - \beta x_{t-1})(y_t - \beta x_t) - \Delta[(x_k - \beta x_{k-1}) - (x_{k+1} - \beta x_k)] \\ b_3 &= \sum_{t=1}^n (y_{t-1} - \beta x_{t-1})^2 + \Delta^2 \end{aligned}$$

On obtient la loi a posteriori conditionnelle de  $\alpha$  suivante

$$\begin{aligned} \pi(\alpha|\beta, \Delta, y, x, x_{n+1}) &\propto [b_1 - 2\alpha b_2 + \alpha^2 b_3]^{-\frac{(n-1)}{2}} \\ &\propto \left[ 1 + \frac{(\alpha - \frac{b_2}{b_3})^2}{n \frac{(\frac{b_1}{b_3} - \frac{b_2^2}{b_3^2})}{n}} \right]^{-\frac{n+1}{2}} \end{aligned}$$

On obtient le noyau d'une loi de student

$$\alpha|y_{n+1}, \beta, \Delta, Y, X, x_{n+1} \sim t_n(\hat{\alpha}, S_2, n)$$

avec

$$\begin{aligned}\hat{\alpha} &= \frac{b_2}{b_3} \\ &= \frac{\sum_{t=1}^n (y_{t-1} - \beta x_{t-1})(y_t - \beta x_t) - \Delta [(x_k - \beta x_{k-1}) - (x_{k+1} - \beta x_k)]}{\sum_{t=1}^n (y_{t-1} - \beta x_{t-1})^2 + \Delta^2}\end{aligned}$$

et

$$\begin{aligned}S_2^2 &= \frac{1}{n} \left( \frac{b_1}{b_3} - \frac{b_2^2}{b_3^2} \right) \\ &= \frac{1}{n} \left( \frac{\sum_{t=1}^n (y_t - \beta x_t)^2 + \Delta^2 - 2\Delta [(y_k - \beta x_k) + (y_{k+1} - \beta x_{k+1})]}{\sum_{t=1}^n (y_{t-1} - \beta x_{t-1})^2 + \Delta^2} \right. \\ &\quad \left. + \frac{(y_{n+1} - \beta x_{n+1} - \alpha(y_n - \beta x_n))^2}{\sum_{t=1}^n (y_{t-1} - \beta x_{t-1})^2 + \Delta^2} \right. \\ &\quad \left. - \frac{\left[ \sum_{t=1}^n (y_{t-1} - \beta x_{t-1})(y_t - \beta x_t) - \Delta [(x_k - \beta x_{k-1}) - (x_{k+1} - \beta x_k)] \right]^2}{\left[ \sum_{t=1}^n (y_{t-1} - \beta x_{t-1})^2 + \Delta^2 \right]^2} \right)\end{aligned}$$

### 2.5.1.3 Loi a posteriori conditionnelle de $\beta$

$$\begin{aligned}\pi(\beta|\alpha, \Delta, y, x, x_{n+1}) &\propto \left[ \sum_{t=1}^n y_t - \alpha y_{t-1} \right]^2 + \Delta^2(1 + \alpha^2) - 2\Delta [(y_k - \alpha x_k) - \alpha(y_{k+1} - \alpha x_{k+1})] \\ &\quad - 2\beta \left[ \sum_{t=1}^n (x_t - \alpha x_{t-1})(y_t + \alpha y_{t-1}) + \Delta(\alpha x_{k-1} - x_k) + \alpha(x_{k+1} - \alpha x_k) \right] \\ &\quad + \beta^2 \sum_{t=1}^n (x_t + \alpha x_{t-1})^2 \\ &\quad + (y_{n+1} - \alpha y_n - \beta(x_{n+1} - \alpha x_n))^2 \Big]^{-\frac{n+1}{2}}\end{aligned}$$

on pose

$$\begin{aligned} c_1 &= \sum_{t=1}^n (y_t - \alpha y_{t-1})^2 + \Delta^2(1 + \alpha^2) - 2\Delta[(y_k - \alpha x_k) - \alpha(y_{k+1} - \alpha x_{k+1})] \\ c_2 &= \sum_{t=1}^n (x_t - \alpha x_{t-1})(y_t + \alpha y_{t-1}) + \Delta(\alpha x_{k-1} - x_k) + \alpha(x_{k+1} - \alpha x_k) \\ c_3 &= \sum_{t=1}^n (x_t + \alpha x_{t-1})^2 \end{aligned}$$

On obtient

$$\begin{aligned} \pi(\beta | \alpha, \Delta, y, x, x_{n+1}) &\propto [c_1 - 2\beta c_2 + \beta^2 c_3]^{-\frac{(n+1)}{2}} \\ &\propto \left[ 1 + \frac{(\beta - \frac{c_2}{c_3})^2}{n \frac{(\frac{c_1}{c_3} - \frac{c_2^2}{c_3^2})}{n}} \right] \end{aligned}$$

On obtient le noyau d'une loi de student

$$\beta | y_{n+1}, \alpha, \Delta, Y, X, x_{n+1} \sim t_n(\hat{\beta}, S_3, n)$$

avec

$$\begin{aligned} \hat{\beta} &= \frac{c_2}{c_3} \\ &= \frac{\sum_{t=1}^n (x_t - \alpha x_{t-1})(y_t + \alpha y_{t-1}) + \Delta(\alpha x_{k-1} - x_k) + \alpha(x_{k+1} - \alpha x_k)}{\sum_{t=1}^n (x_t + \alpha x_{t-1})^2} \\ S_3^2 &= \frac{1}{n} \left( \frac{c_1}{c_3} - \frac{c_2^2}{c_3^2} \right) \\ &= \frac{1}{n} \left( \frac{\sum_{t=1}^n (y_t - \alpha y_{t-1})^2 + \Delta^2(1 + \alpha^2) - 2\Delta[(y_k - \alpha x_k) - \alpha(y_{k+1} - \alpha x_{k+1})]}{\sum_{t=1}^{n+1} (x_t + \alpha x_{t-1})^2} \right. \\ &\quad \left. - \frac{\left( \sum_{t=1}^n (x_t - \alpha x_{t-1})(y_t + \alpha y_{t-1}) + \Delta(\alpha x_{k-1} - x_k) + \alpha(x_{k+1} - \alpha x_k) \right)^2}{\left( \sum_{t=1}^{n+1} (x_t + \alpha x_{t-1})^2 \right)^2} \right) \end{aligned}$$

### 2.5.1.4 Loi a posteriori conditionnelle de $\Delta$

$$\begin{aligned} \pi(\Delta|\alpha, \beta, y, y_{n+1}, x, x_{n+1}) &= \left[ \sum_{t=1}^n (y_t - \beta x_t - \alpha(y_{t-1} - \beta x_{t-1}))^2 \right. \\ &\quad - 2\Delta \left[ (y_k - \beta x_k) - \alpha[(x_k - \beta x_{k-1}) + (y_{k+1} - \beta x_{k+1}) \right. \\ &\quad \left. \left. - \alpha(x_{k+1} - \beta x_k)] \right] + \Delta^2(\alpha^2 + 1) \right. \\ &\quad \left. + (y_{n+1} - \beta x_{n+1} - \alpha(y_n - \beta x_n))^2 \right]^{-\frac{n+1}{2}} \end{aligned}$$

On pose le changement de variables suivant :

$$\begin{aligned} d_1 &= \sum_{t=1}^n (y_t - \beta x_t - \alpha(y_{t-1} - \beta x_{t-1}))^2 + (y_{n+1} - \beta x_{n+1} - \alpha(y_n - \beta x_n))^2 \\ d_2 &= (y_k - \beta x_k) - \alpha[(x_k - \beta x_{k-1}) + (y_{k+1} - \beta x_{k+1}) - \alpha(x_{k+1} - \beta x_k)] \\ d_3 &= \alpha^2 + 1 \end{aligned}$$

On obtient la loi a posteriori conditionnelle de  $\Delta$  suivante :

$$\begin{aligned} \pi(\Delta|\alpha, \beta, y, y_{n+1}, x, x_{n+1}) &\propto [d_1 - 2\Delta d_2 + \Delta^2 d_3]^{-\frac{(n+1)}{2}} \\ &\propto \left[ 1 + \frac{(\Delta - \frac{d_2}{d_3})^2}{\frac{d_1}{d_3} - \frac{d_2^2}{d_3^2}} \right]^{-\frac{n+1}{2}} \end{aligned}$$

On obtient le noyau d'une loi de student

$$\Delta|y, y_{n+1}, \alpha, \beta, x, x_{n+1} \sim t_n(\hat{\Delta}, S_4, n)$$

avec

$$\begin{aligned} \hat{\Delta} &= \frac{d_2}{d_3} \\ &= \frac{(y_k - \beta x_k) - \alpha[(x_k - \beta x_{k-1}) + (y_{k+1} - \beta x_{k+1}) - \alpha(x_{k+1} - \beta x_k)]}{\alpha^2 + 1} \end{aligned}$$

$$\begin{aligned}
 S_4^2 &= \frac{1}{n} \left( \frac{d_1}{d_3} - \frac{d_2^2}{d_3^2} \right) \\
 &= \frac{1}{n} \left( \frac{\sum_{t=1}^{n+1} (y_t - \beta x_t - \alpha(y_{t-1} - \beta x_{t-1}))^2 + (y_{n+1} - \beta x_{n+1} - \alpha(y_n - \beta x_n))^2}{\alpha^2 + 1} \right. \\
 &\quad \left. - \frac{((y_k - \beta x_k) - \alpha[(x_k - \beta x_{k-1}) + (y_{k+1} - \beta x_{k+1}) - \alpha(x_{k+1} - \beta x_k)])^2}{(\alpha^2 + 1)^2} \right)
 \end{aligned}$$

**Remarque 2.1.** De manière similaire, il est possible de calculer la densité prédictive de  $y_{n+2}$  sachant  $y_1, y_2, y_3, \dots, y_n, y_{n+1}$ .

## Chapitre 3

# Étude de simulation

Cette étude de simulation a pour objectif de comprendre comment les valeurs aberrantes additives influencent les estimations des paramètres du modèle de régression à erreurs autocorrélées et les intervalles de crédibilité.

Dans ce chapitre nous allons voir via le langage Matlab l'application de l'échantillonneur de Gibbs pour l'estimation des paramètres d'un modèle de régression simple quand les erreurs sont dépendantes ( $\epsilon_t \sim AR(1)$ ), ainsi que l'estimation d'une observation future, sous contamination AO.

### 3.1 Présentation du modèle

Dans ce qui suit, nous fournissons un exemple de simulation pour étudier la convergence des estimateurs Bayésiens des paramètres  $\beta, \alpha$  et  $\Delta$  à l'aide de l'algorithme de Gibbs. Nous supposons que le processus AR(1) est stationnaire et nous simulons les observations  $y$  en utilisant le modèle suivant

$$\begin{cases} Y_t = 4X_t + \mu_t & t \neq k \\ Y_k = 4X_k + \mu_k + \Delta & t = k \\ \mu_t = 0.5\mu_{t-1} + \epsilon_t \\ \mu_0 = 0.5 & \epsilon_t \sim \mathcal{N}(0, 1) \end{cases} \quad (3.1)$$

Le graphe de la figure suivante nous donne la simulation de  $n = 200, \Delta = 6, k = 100$  observations de la série  $y$  à partir du modèle (3.1)

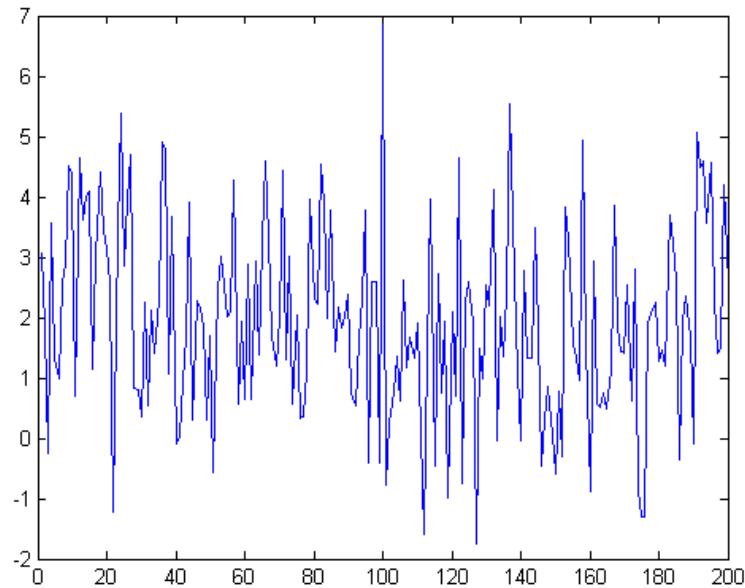


FIGURE 3.1 – Graphe de la série d'observations  $y$  sous contamination AO

## 3.2 Application de l'échantillonneur de Gibbs pour l'estimation des paramètres du modèle sous contamination AO

L'algorithme de Gibbs Sampler est utilisé pour  $N = 500$  itérations afin d'approximer la moyenne a posteriori de  $\theta \in \{\beta, \alpha, \Delta\}$ , en utilisant la fonction de perte quadratique (coût quadratique), définie comme  $\hat{\theta} = \frac{1}{N} \sum_{i=1}^N \theta_i^{(m)}$  avec un "burn-in" période de  $m - 1$  observations pour atténuer l'influence des valeurs initiales, pour  $N$  suites de Gibbs de taille  $m$ .

*l'écart type (std)* et *le rmse* sont donnés comme suit

$$std(\theta) = \sqrt{\frac{1}{N-1} \sum_{t=1}^N (\hat{\theta} - \theta_t^{(m)})^2} \quad (3.2)$$

$$rmse(\theta) = \sqrt{\frac{1}{N} \sum_{t=1}^N (\theta_t^{(m)} - \theta)^2} \quad (3.3)$$

Pour estimer l'intervalle de crédibilité (*IC*) de  $\theta$ , nous ordonnons l'échantillon  $\theta_1^{(m)}, \theta_2^{(m)}, \dots, \theta_N^{(m)}$  de la manière suivante :  $\theta_{(1)}^{(m)}, \theta_{(2)}^{(m)}, \dots, \theta_{(N)}^{(m)}$ . L'intervalle de crédibilité à 95% ( $\alpha = 0.05$ ) de  $\theta$  correspond à l'intervalle délimité par les quantiles  $\frac{\alpha}{2}$ -ième et  $(1 - \frac{\alpha}{2})$ -ième de la distribution a posteriori du paramètre étudié.

On donne ci-dessous les résultats des estimations des paramètres de notre modèle pour différentes tailles d'échantillons .

Paramètres	Vraie-valeurs	$n = 40, (k_1 = 20)$	$n = 200, (k_1 = 100)$	$n = 600, (k_1 = 300)$
$\hat{\beta}_{(rmse)}^{(std)}$	4	$4.0222_{(0.5830)}^{(0.4079)}$	$4.0106_{(0.2496)}^{(0.1782)}$	$4.0011_{(0.1437)}^{(0.1019)}$
$IC(\beta)$		[3.2431, 4.7783]	[3.6509, 4, 3642]	[3.7972, 4.1971]
$\hat{\alpha}_{(rmse)}^{(std)}$	0.5	$0.4645_{(0.2188)}^{(0.1493)}$	$0.4961_{(0.0877)}^{(0.0614)}$	$0.4987_{(0.0135)}^{(0.0094)}$
$IC(\alpha)$		[0.1344, 0.7164]	[0.3715, 0.6060]	[0.4213, 0.5694]
$\hat{\Delta}_{(rmse)}^{(std)}$	6	$6.0163_{(1.4810)}^{(0.9287)}$	$5.9973_{(1.2574)}^{(0.8457)}$	$5.9904_{(1.2411)}^{(0.7223)}$
$IC(\Delta)$		[4.3306, 7.9114]	[4.3284, 7.6048]	[4.5546, 7.3419]

TABLE 3.1 – Résultats des estimations des paramètres  $\beta, \alpha, \Delta$  pour différentes tailles d'échantillons.

### Remarques

- (i) On Remarque que plus la taille de l'échantillon  $n$  augmente plus les estimateurs  $\hat{\alpha}, \hat{\beta}, \hat{\Delta}$  se rapprochent de leur vraie valeurs .
- (ii) On remarque que plus la taille de l'échantillon augmente plus les amplitudes des intervalles de crédibilité de  $\beta, \alpha$  et  $\Delta$  diminuent.
- (iii) On remarque ainsi que nos estimateurs  $\alpha, \beta$  et  $\Delta$  sont convergents puisque le  $rmse$  diminue au même temps que la taille de l'échantillon augmente.

Le tableau suivant donne les résultats des estimations des paramètres de notre modèle pour différentes tailles d'échantillons quand  $\alpha$  est proche de la zone de non stationnarité ( $\alpha = 0.98$ ).

Paramètres	Vraie-valeurs	$n = 40, (k_1 = 20)$	$n = 200, (k_1 = 100)$	$n = 600, (k_1 = 300)$
$\hat{\beta}_{(rmse)}^{(std)}$	4	$4.0465_{(0.6378)}^{(0.3974)}$	$4.0008_{(0.2504)}^{(0.1765)}$	$4.0036_{(0.1426)}^{(0.1006)}$
$IC(\beta)$		[3.2177, 4.8097]	[3.6583, 4.3624]	[3.8046, 4.1932]
$\hat{\alpha}_{(rmse)}^{(std)}$	0.98	$0.9203_{(0.1334)}^{(0.0913)}$	$0.9701_{(0.0291)}^{(0.0205)}$	$0.9770_{(0.0132)}^{(0.0094)}$
$IC(\alpha)$		[0.6721, 1.0245]	[0.9190, 0.9956]	[0.9538, 0.9909]
$\hat{\Delta}_{(rmse)}^{(std)}$	6	$5.9792_{(2.0673)}^{(0.7424)}$	$5.9603_{(1.3878)}^{(0.7279)}$	$6.0164_{(1.2500)}^{(0.7320)}$
$IC(\Delta)$		[4.5229, 7.3575]	[4.4582, 7.3396]	[4.5634, 7.4851]

TABLE 3.2 – Résultats des estimations des paramètres  $\beta, \alpha, \Delta$  pour différentes tailles d'échantillons

### Remarques

- (i) On remarque que tous les estimateurs sont convergents puisque les *rmse* diminuent quand la taille de l'échantillon augmente.
- (ii) On remarque que les amplitudes diminuent quand la taille de l'échantillon augmente.
- (iii) On remarque que les écart-types (*std*) diminuent quand la taille de l'échantillon augmente.

### 3.3 Application de l'échantillonneur de Gibbs pour la prédiction d'une valeur future sous contamination AO

Pour  $n = 200$ ,  $\beta = 4$ ,  $\Delta_{k_1} = 4$ ,  $\sigma = 1$  et différents instants  $k_1$  de contamination. La valeur prédite  $y_{n+1}$  est estimée comme la moyenne a posteriori de la densité prédictive en utilisant l'algorithme de Gibbs sampler pour 500 répétitions. Les résultats de simulation sont données dans la table suivante

Le *rmse* et le *std* sont indiqués comme suit :  $y_{n+1}^{(std)}_{(rmse)}$

$k_1$	la vraie valeur de $y_{n+1}$	$y_{n+1}^{(std)}_{(rmse)}$
$k_1 = 30$	2.0671	2.0993 <sup>(1.1680)</sup> <sub>(1.0678)</sub>
$k_1 = 100$	1.9870	2.0681 <sup>(1.1488)</sup> <sub>(1.0782)</sub>
$k_1 = 199$	2.0058	2.0068 <sup>(1.1588)</sup> <sub>(1.0890)</sub>

TABLE 3.3 – Observation prédite  $y_{n+1}$  sous contamination AO

**Remarque 3.1.** *Pour tout les instants de contamination, on remarque que la valeur estimée de  $y_{n+1}$  est proche de la vraie valeur  $y_{n+1}$ .*

# Conclusion générale et perspectives

Dans ce mémoire, nous avons adopté une approche bayésienne pour estimer les paramètres d'un modèle de régression linéaire qui présente des erreurs autocorrélées et en présence d'une contamination AO.

L'algorithme de Gibbs sampler nous permet d'estimer les paramètres à la fois du modèle de régression linéaire, du modèle autorégressif et l'amplitude de la contamination AO. Pour ce faire, nous avons déterminé les lois a posteriori conditionnelles des paramètres du modèle, ainsi que la densité prédictive d'une observation future.

Nos résultats de simulation ont montré que les paramètres du modèle et la valeur future sont bien estimés.

## Perspectives

- Généraliser l'étude au cas d'un modèle autorégressif d'ordre  $p$ , où  $p \geq 2$ .
- Prendre en compte des distributions a priori informatives.

## Résumé

Au cours de cette étude, nous avons appliqué l'algorithme de Gibbs sampler pour estimer les paramètres d'un modèle de régression linéaire avec des erreurs autocorrélées suivant un  $AR(1)$ , tout en prenant en compte la présence de contamination AO. Nous avons également utilisé cette méthode pour effectuer des prédictions de valeurs futures.

Afin de démontrer l'efficacité de cette méthode, une étude de simulation a été réalisée pour évaluer ses performances.

## Abstract

In this study, we applied the Gibbs sampler algorithm to estimate the parameters of a linear regression model with autocorrelated errors following an  $AR(1)$ , while taking into account the presence of AO contamination. We also used this method to predict future values.

To prove the efficiency of this method, a simulation study was carried out to evaluate its performance.

# Bibliographie

- [1] AIT-MOHAMMED, N. *Approche bayésienne dans la détection des observations aberrantes*. Memoire de magister, 2006.
- [2] AOUIDAD, D. *L'échantillonneur de Gibbs pour l'estimation bayésienne dans un modele de regression lineaire*. Memoire de Master, 2022.
- [3] BERNARDO, J. M., AND SMITH, A. F. *Bayesian theory*, vol. 405. John Wiley & Sons, 2009.
- [4] BERRY, D. A., AND FRISTEDT, B. *London : Chapman and Hall*. Springer, 1985.
- [5] BOX, G. E., JENKINS, G. M., REINSEL, G. C., AND LJUNG, G. M. *Time series analysis : forecasting and control*. John Wiley & Sons, 2015.
- [6] BOX, G. E., AND TIAO, G. C. *A Bayesian approach to some outlier problems*, vol. 55. Oxford University Press, 1968.
- [7] CARLIN, B. P., AND LOUIS, T. A. *Bayes and empirical Bayes methods for data analysis*. Springer, 1997.
- [8] CASELLA, G., AND GEORGE, E. I. *Explaining the Gibbs sampler*, vol. 46. Taylor & Francis, 1992.
- [9] CHALONER, K., AND BRANT, R. *A Bayesian approach to outlier detection and residual analysis*, vol. 75. Oxford University Press, 1988.
- [10] CHEN, C. W. *Detection of additive outliers in bilinear time series*, vol. 24. Elsevier, 1997.
- [11] FREEDMAN, L., AND SPIEGELHALTER, D. *The assessment of subjective opinion and its use in relation to stopping rules for clinical trials*, vol. 32. Wiley Online Library, 1983.

- 
- [12] GALEANO, P., AND PEÑA, D. *Finding outliers in linear and nonlinear time series*. Springer, 2013.
- [13] GAMERMAN, D., AND LOPES, H. F. *Markov chain Monte Carlo : stochastic simulation for Bayesian inference*. CRC press, 2006.
- [14] GELMAN, A., CARLIN, J. B., STERN, H. S., AND RUBIN, D. B. *Bayesian data analysis*. Chapman and Hall, 1995.
- [15] GEMAN, S., AND GEMAN, D. *Stochastic relaxation, Gibbs distributions, and the Bayesian restoration of images*. No. 6. IEEE, 1984.
- [16] GILKS, W. R., RICHARDSON, S., AND SPIEGELHALTER, D. *Markov chain Monte Carlo in practice*. CRC press, 1995.
- [17] HAMILTON, J. D. *Time series analysis*. Princeton university press, 2020.
- [18] JACK, G., AND IRWIN, G. *A Bayesian approach to prediction in the presence of spurious observations for several models*, vol. 13. Taylor & Francis, 1984.
- [19] JUDGE, G., HILL, R., GRIFFITHS, W., LUTKEPOHL, H., AND LEE, T. *Introduction to the Theory and Practice of Econometrics. 2d ed. New York : John Wiley & Sons*. 1988.
- [20] ROBERT, C. *Le choix bayésien : Principes et pratique*. Springer Science & Business Media, 2006.
- [21] ROBERT, C. P. *L'analyse statistique bayésienne*. Economica, 1992.
- [22] ROBERT, C. P., CASELLA, G., AND CASELLA, G. *Monte Carlo statistical methods*, vol. 2. Springer, 1999.
- [23] ROBERT, C. P., CASELLA, G., AND CASELLA, G. *Introducing monte carlo methods with r*, vol. 18. Springer, 2010.
- [24] ROBERTS, G. O., AND SMITH, A. F. *Simple conditions for the convergence of the Gibbs sampler and Metropolis-Hastings algorithms*, vol. 49. Elsevier, 1993.
- [25] TIAO, G. C., AND ZELLNER, A. *On the Bayesian estimation of multivariate regression*, vol. 26. Wiley Online Library, 1964.
- [26] ZELLNER, A. *Bayesian and non-Bayesian analysis of the log-normal distribution and log-normal regression*, vol. 66. Taylor & Francis, 1971.