

Remerciements

Je tiens à remercier en premier lieu le Dieu le tout puissant pour la persévérance, la santé, l'ouverture d'esprit et l'inspiration dont il ma a gratifié.

Aussi, le présent mémoire n'aurait pas vu le jour sans l'aide bienveillante, la compréhension et le dévouement d'un certain nombre de personnes à qui j'adresse mon vif remerciements pour :

Ma promotrice **Mme BELKACEM Cherifa** pour sa grande disponibilité et d'avoir accepté de diriger ce travail.

Mes vifs remerciements vont également au Professeur **M.Fellag Hocine** pour tous ce qui nous a appris et d'avoir accepté de juger ce travail.

Je tiens aussi à exprimer mes remerciements à **Mme MERABET Dalila**, Maitre de conférences au département de mathématiques, pour avoir accepté de participer à ce jury.

Je remercie sincèrement tous mes enseignants de proba.stat.

Je remercie mes amis de la promotion 2018-2019 probabilité et statistique.

Table des matières

Table des matières	1
Introduction générale	5
1 Généralités sur les séries chronologiques	8
1.1 Définition et domaines d'applications d'une série chronologique	8
1.2 Composantes d'une série chronologique	11
1.3 Description schématique de l'étude complète d'une série chronologique . . .	11
1.4 Modélisation d'une série chronologique	12
1.4.1 Modèle additif :	13
1.4.2 Modèle multiplicatif :	13
1.4.3 Modèle mixte :	13
1.5 Opérateurs définis sur une série chronologique	14
1.5.1 Opérateur retard	14
1.5.2 Opérateur de différence	14
1.6 Stationnarité	15
1.7 propriétés	16
1.7.1 Elimination du trend en l'absence de saisonnalité	16
1.7.2 Elimination du trend et de la saisonnalité	18
1.8 Processus stochastique	19
1.8.1 La fonction d'autocovariance et d'autocorrélation	19
1.9 Bruit blanc	20

1.10	Modélisation ARIMA des séries chronologiques	21
1.10.1	Processus autorégressif AR(p)	21
1.10.2	Processus à moyenne mobile MA(q)	21
1.10.3	Processus ARMA(p,q)	23
1.10.4	Le processus intégrée I(d)	24
1.10.5	Le processus ARIMA(p,d,q)	24
2	Les modèles de rupture	26
2.1	Introduction	26
2.2	Les différents modèles de rupture	27
2.2.1	Rupture unique	27
2.2.2	Ruptures multiples	28
2.2.3	Rupture épidémique	28
2.3	Tests statistiques de rupture	29
2.3.1	Tests de ruptures multiples	29
2.3.2	Tests de rupture unique à une position donnée	31
2.3.3	Exemple de rupture dans la moyenne dans le cas des données gaussiennes	32
2.4	Approche Bayésienne dans les modèles de rupture	35
2.4.1	Modèle de rupture dans le cas où les variables sont indépendantes	36
2.4.2	Modèle de rupture dans le cas où les variables sont dépendantes	37
2.4.2.1	Modèle avec des erreurs autocorrélées	38
2.4.2.2	Modèle de régression linéaire	39
3	Estimation Bayésienne d'un point de rupture dans un processus autorégressif	44
3.1	Introduction	44
3.2	Présentation du modèle	45
3.3	Estimation Bayésienne	45

3.3.1	La fonction de vraisemblance :	46
3.3.2	Les lois a priori	48
3.3.3	la densité a posteriori jointe	49
3.3.4	Les densités a posteriori marginales	52
3.4	Estimateur Bayésien du point de rupture sous différentes fonctions coûts .	53
3.4.1	La fonction coût quadratique	53
3.4.2	La fonction coût absolu	54
3.4.3	La fonction coût 0-1	54
3.5	Analyse Bayésienne en utilisant des lois a priori non informatives	54
3.5.1	la distribution a priori jointe	55
3.5.2	La distribution a posteriori jointe de β_1 , β_2 et de m	55
3.6	Simulation	57
 Conclusion et perspectives		 70
 Bibliographie		 70

Introduction générale

Le monde tel qu'on le connaît est rempli de phénomènes dépendant du temps. En statistique, l'un des modèles fréquemment utilisé est celui des séries chronologiques (temporelles).

Les séries chronologiques ont commencé à être utilisées en météorologie dans l'étude des températures et de la pluviométrie. L'étude descriptive de ces séries n'était pas suffisante pour prévoir leurs comportements futurs. Le besoin de prévision de ce type de séries a emmené les scientifiques à construire des modèles mathématiques aptes à répondre à ce besoin. La majeure partie de la méthodologie statistique classique repose sur des séries d'observations indépendantes. Généralement, ce manque d'indépendance est considéré comme un handicap. Ainsi, un des objectifs d'une bonne expérimentation est d'éliminer cette dépendance. Toutefois, dans le cas de l'analyse des séries chronologiques nous nous occupons de données qui se développent dans le temps et telle que chaque observation peut dépendre dans une certaine mesure des observations précédentes. C'est sur cette dépendance que se focalise l'étude des séries chronologiques.

Lorsqu'un changement brusque apparaît dans une série chronologique, nous nous intéressons à un modèle de rupture (exemple de réchauffement climatique dans l'étude des températures).

L'étude de rupture de modèles est un passage incontournable pour les économètres et les statisticiens. En effet la stabilité d'un modèle si elle venait à faire défaut doit être décelée sinon elle risque de conduire à des résultats erronés et entraîner des interprétations fausses.

Le problème de rupture est un problème qui surgissent dans beaucoup de domaines, tel que l'économie, le traitement du signal, la biologie, la médecine, l'agronomie, la météorologie, etc... .

Parmi les chercheurs qui se sont intéressés au problème de rupture de modèle, Page (1954,1955,1957) est considéré comme le pionier dans ce domaine. En s'intéressant au contrôle de qualité, il a mis au point un test non paramétrique basé sur les sommes cumulées (CUSUMS test) pour déceler une éventuelle existence d'un changement dans la moyenne d'une suite X_1, \dots, X_n de variables aléatoires indépendantes . Chernoff et Zacks (1964) ont étudié une suite de variables aléatoires gaussiennes et indépendantes et ont proposé un test (paramétrique) Bayésien pour mettre en évidence un changement dans la moyenne. A la fin des années soixantes et début des années soixante dix, beaucoup de travaux ont été initiés pour des modèles de régression à deux régimes où la méthodologie Bayésienne a été largement sollicitée; on peut cité entre autres les travaux de Ferreira (1975) qui a déterminé les ditributions a posteriori des paramètres de régression dans le cas d'un modèles de rupture à deux phases (modèle où n'apparait qu'une seule rupture). Abdelli.Z (1993) a considéré le problème de rupture dans les séries chronologiques (dans un AR(p), $p \geq 1$). Plotr Kokoszka et al (2000) se sont intéressés au modèle ARCH avec rupture. Fotopoulos et al (2010), Fearnhead (2006) et Rigaiil et al (2012) se sont intéressés au problème de ruptures multiples.

Le document comporte trois chapitres :

Le premier chapitre comprend quelques généralités sur les série chronologiques et leurs modélisations .

Dans le deuxième chapitre nous abordons les modèles de rupture, nous donnons quelques méthodes de détection de rupture (tests d'hypothèses, estimation Bayésienne). Nous citons également quelques exemples de modèles ayant subi un changement de

structure à un instant donné.

Dans le troisième chapitre nous nous intéressons à l'estimation d'un point de rupture dans un processus autorégressif d'ordre un, en utilisant l'estimation Bayésienne.

Chapitre 1

Généralités sur les séries chronologiques

La théorie des séries chronologiques (ou temporelles) est l'évolution au cours du temps d'un phénomène dans le but de décrire, expliquer puis prévoir ce phénomène dans le futur. Nous disposons ainsi d'observations à des dates différentes, c'est à dire d'une suite de valeurs numériques indicées par le temps.

Les séries temporelles sont utilisées en astronomie 1906, en météorologie 1968, en biologie 1960 et en économie 1971, écologie, médecine...etc.

1.1 Définition et domaines d'applications d'une série chronologique

Nous appelons "série chronologique" la suite d'observations d'une famille de variables aléatoires réelles notées $(X_t)_{t \in \theta}$ (ou $X_t, t \in \theta$).

Où l'ensemble θ est l'espace des temps qui peut être :

- **Discret** : dans ce cas $\theta \subset \mathbb{Z}$. Les dates d'observations sont le plus souvent équidistantes.

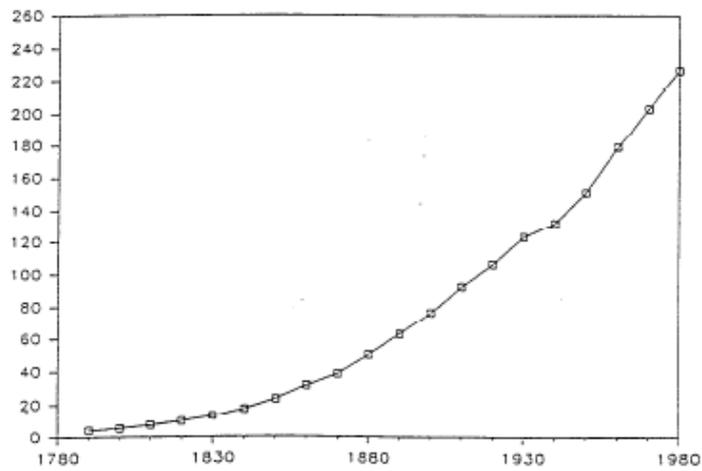
Exemple : le nombre de passagers dans les lignes aériennes.

- **Continu** : dans ce cas $\theta \subset \mathbb{R}$. On dispose d'une infinité d'observations issues d'un

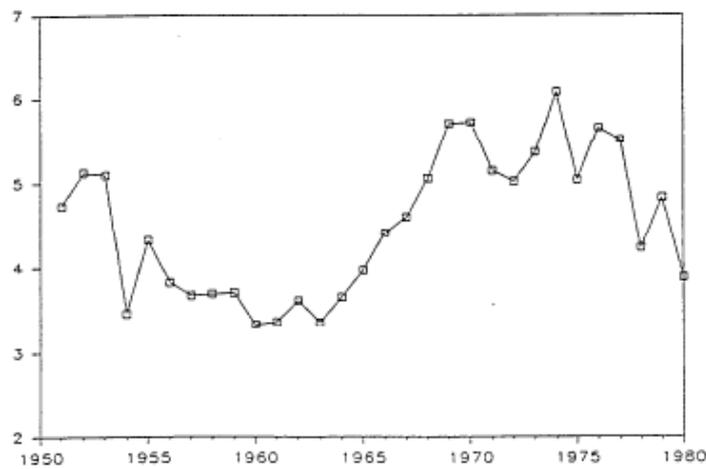
processus $(X_t)_{t \in \theta}$.

Exemple : résultat d'un électrocardiogramme.

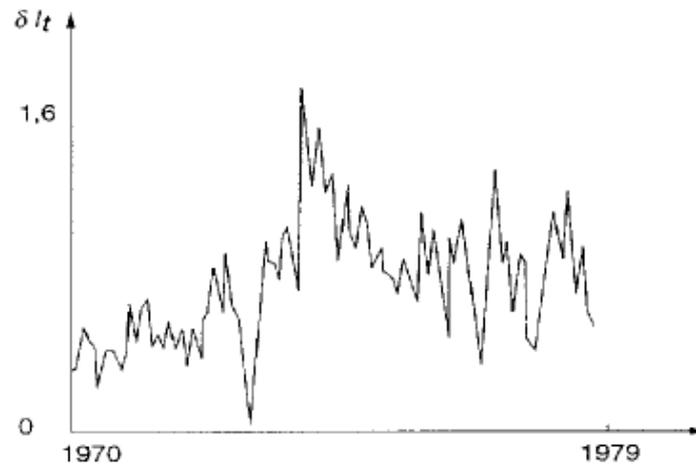
Exemples des séries chronologiques



(a) Population des Etats-Unis, 1790–1980



(b) Nombre de grèves aux Etats-Unis, 1951-1980



(c) **Accroissement relatif mensuel
de l'indice des prix**

Les domaines d'applications d'une série chronologique :

La théorie des séries temporelles (ou chronologiques) est appliquée de nos jours dans des domaines aussi divers que :

1. Finance et économétrie : évolution des indices boursiers, des prix, des données économiques des entreprises, des ventes et achats, des productions agricoles ou industrielles.
2. Médecine : suivi des évolutions des pathologies.
3. Démographie : analyse de l'évolution d'une population.
4. Météorologie : analyse de données climatiques.
5. Traitement d'images : analyse d'images satellites.
6. Hydrologie : moyenne de la concentration en nitrates dans les eaux superficielles.
7. Ecologie, Environnement, Traitement du signal, Traitement des données, etc.

1.2 Composantes d'une série chronologique

La série chronologique (X_t) est la résultante de différentes composantes fondamentales :

- **La tendance (ou trend)** (Z_t) : représente l'évolution à long terme de la série étudiée. Elle traduit le comportement "moyen" de la série. (Il peut être aléatoire comme il peut être déterministe)
- **La composante saisonnière (ou saisonnalité)** (S_t) : correspond à un phénomène qui se répète à des intervalles de temps réguliers (périodiques). En général c'est un phénomène saisonnier d'où le terme de variations saisonnières. (Déterministe)
- **La composante résiduelle (ou bruit ou résidu)** (ϵ_t) : correspond à des fluctuations irrégulières, en général de faible intensité mais de nature aléatoire.

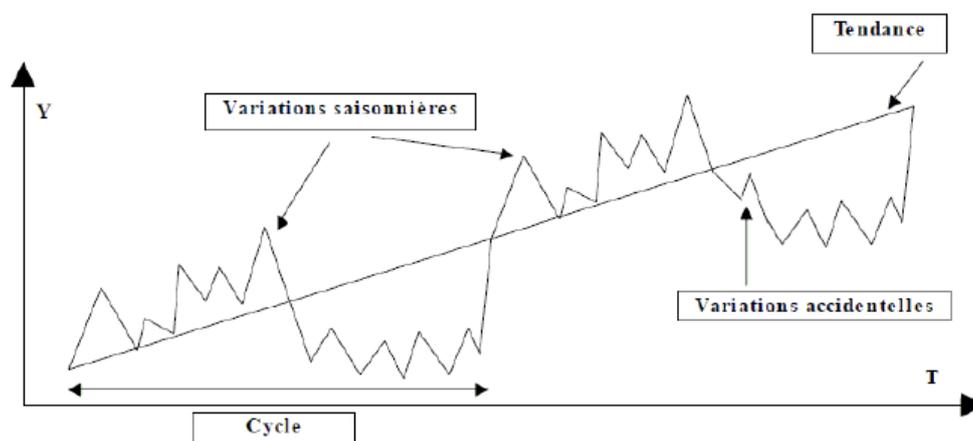


FIG. 1.2 – Les composantes d'une série chronologique

1.3 Description schématique de l'étude complète d'une série chronologique

L'un des objectifs principaux de l'étude d'une série chronologique est la prévision des valeurs futures de cette série. Pour cela, on a besoin de connaître ou tout au moins de

modéliser le mécanisme de production de la série chronologique.

Notons que les variables X_t , ne sont pas souvent indépendantes (on peut s'attendre en effet à ce que des observations relativement proches dans le temps soient liées), ni identiquement distribuées (dans la plupart des cas, le phénomène évolue, se modifie au cours du temps ce qui entraîne que les variables le décrivant ne sont pas équidistribuées). Cela nécessite des méthodes statistiques de traitement et de modélisation spécifiques, puisqu'en particulier dans un cadre standard (celui de la description d'un échantillon) les méthodes statistiques classiques sont basées sur des hypothèses d'indépendance. Schématiquement, les principales étapes de traitement d'une série chronologique sont les suivantes :

1. correction des données
2. observation de la série
3. modélisation (avec un nombre fini de paramètres)
4. analyse de la série à partir de ses composantes
5. diagnostic du modèle - ajustement au modèle
6. prédiction (= prévision)

Un autre objectif de l'analyse intéressant est **la détection de ruptures** résultantes, par exemple, d'un changement de politique (économique).

La prévision de ces dates de rupture est bien évidemment très importante.

1.4 Modélisation d'une série chronologique

Un modèle est l'image simplifiée de la réalité. Un modèle peut être meilleur qu'un autre pour décrire la réalité. Donc la question qui se pose : Comment diagnostiquer un modèle ?

Les modèles déterministes relèvent de la statistique descriptive . Il ne font intervenir que de manière implicite le calcul des probabilités et consistent à supposer que l'observation de

la série à la dates t est une fonction du temps t et d'une variable ϵ_t centrée faisant office d'erreur au modèle, représentant la différence entre la réalité et le modèle proposé :

$$X_t = f(t, \epsilon_t)$$

On suppose de plus que les ϵ_t sont décorrélées.

les modèles de ce type les plus utilisés sont les suivants :

1.4.1 Modèle additif :

C'est le "modèle classique de décomposition" dans le traitement des modèles d'ajustement. La variable X_t s'écrit comme la somme de trois termes :

$$X_t = Z_t + S_t + \epsilon_t$$

Où Z_t représente la tendance, S_t la saisonnalité (la partie déterministe) et ϵ_t les composantes "erreurs au modèle" qui sont aléatoires, indépendantes et identiquement distribuées.

1.4.2 Modèle multiplicatif :

Certaines séries temporelles se comportent selon un modèle multiplicatif. On a alors :

$$X_t = Z_t S_t \epsilon_t \quad \text{pour tout } t \in \{1, \dots, n\}$$

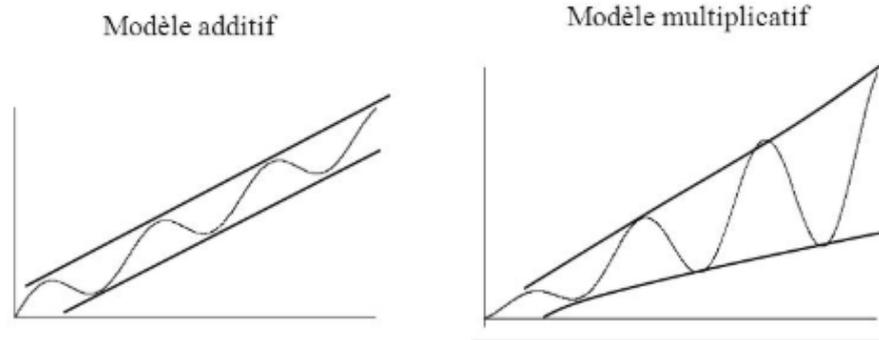
On remarque qu'on se ramène naturellement à un modèle additif par passage au log :

$$\log(X_t) = \log(Z_t) + \log(S_t) + \log(\epsilon_t) \quad \text{pour tout } t \in \{1, \dots, n\}$$

1.4.3 Modèle mixte :

Il s'agit là de modèles où addition et multiplication sont utilisées. Nous pouvons supposer par exemple que la composante saisonnière agit de façon multiplicative alors que les fluctuations irrégulières sont additives :

$$X_t = Z_t S_t + \epsilon_t \quad \text{pour tout } t \in \{1, \dots, T\}$$



1.5 Opérateurs définis sur une série chronologique

1.5.1 Opérateur retard

C'est un opérateur qui transforme une variable X_t en sa valeur passée. Il est désignée par la lettre B (encore noté L) et tel que :

$$\left\{ \begin{array}{l} BX_t = X_{t-1} \text{ (processus décalé d'une unite de temps)} \\ B^2 X_t = B(BX_t) = BY_{t-1} = X_{t-2} \text{ où } Y_{t-1} = BX_t. \\ \cdot \\ \cdot \\ \cdot \\ B^d X_t = X_{t-d} \text{ (processus décalé de } d \text{ unites de temps)}. \end{array} \right.$$

1.5.2 Opérateur de différence

L'opérateur ∇ fait la différence entre le processus et sa version décalée d'une unité de temps. Cet opérateur se construit en utilisant l'opérateur retard.

$$\nabla X_t = X_t - X_{t-1} = X_t - BX_t = (1 - B)X_t$$

et donc on à l'équivalence $\nabla = 1 - B$ en termes d'opérateurs.

Cependant, nous pouvons facilement les généraliser à l'ordre d quelconque. Ainsi, appliqué à la série chronologique X_t , pour tout $d \geq 1$, on note :

$$B^d = X_{t-d} \text{ (l'opérateur retard)}$$

De même,

$$\nabla^d X_t = (1 - B)^d X_t \text{ et } \nabla_d X_t = (1 - B^d) X_t = X_t - X_{t-d}$$

Par exemple lorsque $d = 2$, on obtient

$$B^2 X_t = X_{t-2} \text{ et } \nabla_2 X_t = (1 - B^2) X_t = X_t - X_{t-2}$$

et

$$\begin{aligned} \nabla^2 X_t &= (1 - B)^2 X_t = (1 - 2B + B^2) X_t \\ &= X_t - 2B X_t + B^2 X_t \\ &= X_t - 2X_{t-1} + X_{t-2} \\ &= (X_t - X_{t-1}) - (X_{t-1} - X_{t-2}) \\ &= \nabla(\nabla X_t) \end{aligned}$$

1.6 Stationnarité

La notion de stationnarité caractérise la capacité d'un processus à se décorrélérer totalement de l'indice temporel. Ainsi la loi du processus, bien que restant souvent inconnue, sera bien plus aisée à manipuler par l'intermédiaire de ses propriétés d'espérance et de covariance.

Définition 1.1. *stationnarité forte (stricte)* : Une série chronologique est dite strictement (fortement) stationnaire si $\forall n \geq 0, \forall t_0, t_1, \dots, t_n, h \in T$, $(X_{t_0}, X_{t_1}, \dots, X_{t_n})$ est de même loi que $(X_{t_0+h}, X_{t_1+h}, \dots, X_{t_n+h})$

En particulier X_t et $X_{t+h} = X_s$ ont la même loi $\forall t, s \in T$

C'est à dire la loi de X_t est la même que celle de X_s .

Cette définition illustre bien le fait que l'indice temporel ne joue plus aucun rôle dans le comportement de la série chronologique puisque, en considérant tous les vecteurs de même taille translatés dans le temps, la loi du processus reste inchangée. La loi d'un processus strictement stationnaire ne dépend donc pas du temps. En pratique, il est nécessaire d'être en mesure d'évaluer la loi jointe du processus, ce qui peut s'avérer très difficile. C'est pourquoi l'on introduit généralement une stationnarité plus faible et moins contraignante.

Définition 1.2. Stationnarité faible : une série chronologique $X_t, t \in \mathbb{Z}$ est dite faiblement stationnaire (ou "stationnaire au second ordre" ou même simplement "stationnaire") si :

1. $E(X_t) = m, \quad \forall t \in \mathbb{Z}$ l'espérance mathématique est constante,
2. $V(X_t) = \sigma^2, \quad \forall t \in \mathbb{Z}$ la variance est constante,
3. $cov(X_t, X_{t-h}) = \Gamma_h, \quad \forall h, t \in \mathbb{Z}$ (indépendant de t).

Exemples :

Exemple 01 : Un exemple de processus stationnaire très utilisé est le bruit blanc que nous allons voir par la suite. Comme le bruit blanc : $E(\epsilon_t) = 0, \quad var(\epsilon_t) = \sigma^2$, et $cov(\epsilon_t, \epsilon_s) = 0 \quad si \quad s \neq t$ Donc c'est un processus stationnaire.

Exemple 02 : Un processus construit à partir d'une série temporelle qui présente une tendance ou une saisonnalité n'est pas stationnaire (la moyenne dépend du temps t).

1.7 propriétés

En pratique, les séries temporelles résiduelles ne sont pas nécessairement stationnaires. Un prétraitement est alors nécessaire pour supprimer la tendance et la saisonnalité d'une part comme usuellement mais aussi pour "stationnariser" la série résiduelle. Donc on s'intéresse à éliminer la partie déterministe pour faire l'étude sur la partie aléatoire.

1.7.1 Elimination du trend en l'absence de saisonnalité

Il faudra éliminer la tendance (ou trend) représentant l'évolution à moyen terme du phénomène étudié. Cette tendance agit comme une forte corrélation entre les variables X_t qui n'exprime aucune liaison à caractère explicatif. Nous verrons comment enlever cette tendance pour voir si de telles liaisons existent afin d'étudier que les corrélations sans tendance qui sont les quantités qui nous intéressent pour expliquer un phénomène observé.

Supposons, dans un premier temps, que la partie déterministe du modèle soit uniquement composée d'un trend Z_t :

$$X_t = Z_t + \epsilon_t$$

Dans ce cas il y'a plusieurs méthodes pour estimer Z_t est parmi ces méthodes nous avons :

Méthode des différences itérées

$$X_t = Z_t + \epsilon_t \quad (1.1)$$

Supposons à présent que l'équation de tendance, soit de la forme d'un polynôme d'ordre k :

$$Z_t = \sum_{j=0}^k a_j t^j$$

Si nous appliquons à Z_t l'opérateur de différence ∇ (d'ordre 1), on obtient :

$$\begin{aligned} \nabla Z_t &= Z_t - Z_{t-1} \\ &= \sum_{j=0}^k a_j t^j - \sum_{j=0}^k a_j (t-1)^j \end{aligned}$$

En développant $(t-1)^j$, nous obtenons, dans le membre de droite :

$$\sum_{j=0}^k a_j (t-1)^j = \sum_{j=0}^k a_j (t)^j + \sum_{j=1}^k b_{j-1} (t)^{j-1}$$

Pour certains coefficients b_t donc, ∇Z_t devient :

$$\begin{aligned} \nabla Z_t &= \sum_{j=1}^k b_{j-1} (t)^{j-1} \\ &= \sum_{i=0}^{k-1} b_i (t)^i \quad \text{avec } i = j - 1 \end{aligned}$$

Et nous voyons que l'opérateur ∇ a diminué l'ordre du polynôme Z_t : Z_t est d'ordre k , tandis que ∇Z_t est d'ordre $k - 1$. Dès lors, si on applique k fois l'opérateur ∇ au polynôme Z_t , nous obtiendrons :

$$\nabla^k Z_t = k! a_k = \text{constante indépendante de } t$$

et donc, appliqué à (1.1), l'opérateur ∇^k livre :

$$\nabla^k X_t = \text{constante} + \nabla^k \epsilon_t.$$

Bien entendu, en pratique, nous ignorons le degré du polynôme Z_t . C'est pourquoi nous appliquons l'opérateur ∇ à la série observée autant de fois qu'il est nécessaire pour obtenir une série $\nabla^k X_t$ dont la moyenne est nulle (et donc seule la composante aléatoire est présente). En pratique k est assez petit (de l'ordre de 1,2,3).

1.7.2 Élimination du trend et de la saisonnalité

Dans le même ordre d'idée, nous corrigerons les éventuelles variations saisonnières qui résultent d'un comportement cyclique dans la série observée. Nous verrons comment éliminer cette saisonnalité afin de nous intéresser uniquement aux composantes aléatoires de la série chronologique.

Supposons à présent que, en plus d'un trend Z_t , la série comporte une composante saisonnière S_t . On considère donc le modèle

$$X_t = Z_t + S_t + \epsilon_t \quad (1.2)$$

où comme nous l'avons spécifié plus haut,

- S_t est une fonction périodique en t , $S_t = S_{t+d}$, pour un certain $d > 0$.
- S_t ne comporte pas de composante "trend", et donc :

$$\sum_{j=1}^d S_{t+j} = 0$$

Méthode des différences On peut également appliquer la méthode des différences pour éliminer à la fois le trend et la saisonnalité. Etant donné le modèle (1.2), appliquons l'opérateur de différence d'ordre d , ∇_d , où d est la période de S_t à la série observée X_t :

$$\begin{aligned}\nabla_d X_t &= X_t - X_{t-d} \\ &= Z_t + S_t + \epsilon_t - Z_{t-d} - S_{t-d} - \epsilon_{t-d} \\ &= (Z_t - Z_{t-d}) + (S_t - S_{t-d}) + (\epsilon_t - \epsilon_{t-d})\end{aligned}$$

Où, par définition, le terme $S_t - S_{t-d}$ est nul. ∇_d est donc une série chronologique désaisonnalisée, mais contenant un trend $\nabla_d Z_t$. Les méthodes d'élimination du trend en l'absence de saisonnalité s'appliquent donc.

1.8 Processus stochastique

Nous parlons de processus stochastique, dit aussi aléatoire, lorsque l'évolution d'une variable dans le temps est imprévisible, c'est à dire qu'il est impossible, connaissant la position de la variable au temps t de prédire avec exactitude sa position au temps $t + \Delta_t$: Un processus non stochastique est dit déterministe, c'est à dire que l'on peut rendre compte au moyen d'une fonction du temps [$x = f(t)$] capturant l'ensemble de la variance de la serie. Un processus stochastique comprend généralement une partie déterministe, auquel s'ajoute une partie stochastique également appelé bruit blanc.

De ce fait, une série chronologique est une réalisation d'un processus stochastique.

1.8.1 La fonction d'autocovariance et d'autocorrélation

Définition 1.3. La fonction d'autocovariance Soit $X_t, t \in T$ un processus avec $var(X_t) < \infty, \forall t \in T$ Alors la fonction d'autocovariance γ_h de X_t est :

$$\gamma_X(r, s) = cov(X_r, X_s) = E[(X_r - E(X_r))(X_s - E(X_s))] = E(X_r, X_s) - E(X_r)E(X_s). \quad r, s \in T$$

Définition 1.4. Fonction d'autocorrélation

Soit $X = X_1, \dots, X_n$; un processus à temps discret. La fonction de corrélation (ou d'autocorrélation) de X est donnée par :

$$p(t, t + h) = \frac{\text{cov}(X_t, X_{t+h})}{\sqrt{\text{Var}(X_t)\text{Var}(X_{t+h})}}$$

Nous pouvons ainsi considérer l'autocorrélation qui est la covariance corrigée de la variance.

1.9 Bruit blanc

On appelle bruit toute variable imprévisible d'une quantité dans le temps. Le bruit blanc est le plus aléatoire des bruits. dans ce cas, il n'existe aucune corrélation entre les accroissements successifs de la quantité. Autrement dit la fonction d'auto-corrélation est donc nulle.

Définition 1.5. Un processus $(\epsilon_t, t \in \mathbb{Z})$ est dit **bruit blanc** s'il :

- Sont de même moyenne $E(\epsilon_t) = 0, \forall t$
- De même variance $\text{var}(\epsilon_t) = \sigma^2, \forall t$
- Et non corrélées.

Dans le cas de processus gaussiens. On parle alors de bruit blanc gaussien.

Le bruit blanc, même non portant de la structure de corrélation non triviale ("processus purement aléatoire"), jouera un rôle prépondérant dans notre modélisation : Il intervient comme module de base pour construire des processus stationnaires portant des corrélations intéressantes ainsi qu'il servira comme modèle pour les "résidus" d'une modélisation - la composante qui reste après avoir modélisé toute structure de corrélation intéressante et qui est supposée de ne plus contribuer à cette structure non triviale. En effet, tester si ces résidus sont compatibles avec un bruit blanc sera un des moyens principaux pour une diagnostique du modèle ajusté aux données stationnaires mais autocorrélées.

Cependant, les résidus ou erreurs peuvent être modélisés autrement que par des bruits blancs (exemple : processus,...).

1.10 Modélisation ARIMA des séries chronologiques

La classe de modèle ARIMA est un outil de prévision important, il signifie "Moyen-mobile Intégrée Autorégressif"

1.10.1 Processus autorégressif AR(p)

Les premiers processus autorégressifs ont été introduits par George Udny Yule dans ([a method of investigating periodicity in distributed series]). Dans cet article, Yule utilise le premier modèle autorégressif pour modéliser la série chronologique du nombre de taches solaires plutôt que la méthode du périodogramme de Schuster ([the Periodicity of Sun-spots]). Un processus autorégressif est un processus où l'on écrit une observation au temps t comme une combinaison linéaire des observations passées plus un certain bruit blanc.

Définition 1.6 (AR(p)). *La suite $X_t : t > 0$ est un processus autoregressif d'ordre p ($p > 0$) s'il peut s'écrire sous la forme suivante :*

$$X_t = \sum_{k=1}^p \phi_k X_{t-k} + \epsilon_t \quad \text{ou} \quad \epsilon_t \sim N(0, \sigma_\epsilon^2)$$

Les ϕ_k ($k = 1, 2, 3, \dots, p$) constituent les paramètres du modèle.

Dans ce cas, on note $X_t \sim AR(p)$, de la même façon, on peut réécrire un processus $AR(p)$ avec un polynôme $\phi(L)$ qui multipliera X_t cette fois-ci :

$$\phi(L)X_t = \epsilon_t \quad \text{avec} \quad \phi(L) = 1 - \phi_1 L - \phi_2 L^2 - \dots - \phi_p L^p$$

1.10.2 Processus à moyenne mobile MA(q)

C'est Eugen Slutsky qui, en 1927, a introduit pour la première fois les processus à moyenne mobile. La définition suivante présente ce processus.

Définition 1.7 (MA(q)). On dit que la suite $X_t : t \geq 0$ est un processus à moyenne mobile d'ordre q , $q > 0$ si celui-ci peut s'écrire sous la forme suivante :

$$X_t = \sum_{k=1}^q \theta_k \epsilon_{t-k} + \epsilon_t \quad \text{où} \quad \epsilon_t \sim N(0, \sigma_\epsilon^2)$$

où les $\theta_k (k = 1, 2, \dots, q)$ sont les paramètres du modèle.

Dans ce cas, on note $X_t \sim MA(q)$.

La causalité : Un processus est dit **causal** s'il existe une suite a_k réelle telle que $\sum_{k=0}^{\infty} |a_k| < \infty$ et que

$$X_t = \sum_{k=0}^{\infty} a_k \epsilon_{t-k} \quad \text{où} \quad \epsilon_t \sim N(0, \sigma_\epsilon^2)$$

Parfois, lorsque l'on parle d'un processus causal, on dit que celui-ci a une représentation $MA(\infty)$.

Remarque : Tout processus $MA(q)$ est causal.

L'inversibilité : Un processus est dit **inversible** s'il existe une suite b_k réelle telle que $\sum_{k=0}^{\infty} |b_k| < \infty$ et

$$\epsilon_t = \sum_{k=0}^{\infty} b_k X_{t-k} \quad \epsilon_t \text{ est un bruit blanc}$$

Une autre façon de dire qu'un processus est inversible est de dire qu'il possède une représentation $AR(\infty)$.

Remarque : Avec cette définition, tout processus $AR(p)$ est inversible.

Théorème 1.1 (Peter J.Brockwell and Richard A.Davis (1991)). Un processus autoregressif $AR(p)$ est causal et stationnaire si et seulement si son polynôme $\phi(z)$ est tel

que

$$\phi(z) \neq 0 \text{ pour toute } z \in \mathbb{C} \text{ tel que } |z| \leq 1$$

En d'autres mots, toutes les racines de $\phi(z)$ sont de norme plus grande que 1.

Théorème 1.2 (Peter J. Brockwell and Richard A. Davis (1991)). *Un processus à moyenne mobile MA(q) est inversible si et seulement si son polynôme $\theta(z)$ est tel que*

$$\theta(z) \neq 0 \text{ pour toute } z \in \mathbb{C} \text{ tel que } |z| \leq 1$$

On note la ressemblance de cet énoncé avec le théorème de stationnarité et de causalité pour les processus autorégressifs.

1.10.3 Processus ARMA(p,q)

Nous pouvons à présent introduire une classe de processus linéaires qui est très importante dans la modélisation des séries chronologiques : les processus ARMA (Auto-Regressive Moving Average). Ces processus sont linéaires et, sous certaines conditions, stationnaires. Leur linéarité fournira une théorie simple de la prévision.

Herman Wold (1938) a montré que les processus ARMA pouvaient être utilisés pour modéliser n'importe quelle série stationnaire pour autant que les ordres p et q soient bien choisis.

Définition 1.8 (ARMA(p,q)). *Un processus X_t , où $t \in \mathbb{Z}$ est dit ARMA(p,q) de moyenne nulle si :*

$$\left\{ \begin{array}{l} (X_t) \text{ est stationnaire} \\ X_t - \sum_{k=1}^p \phi_k X_{t-k} = \epsilon_t + \sum_{j=1}^q \theta_j \epsilon_{t-j}. \end{array} \right.$$

Une représentation équivalente et plus concise est la suivante :

$$\phi(L)X_t = \theta(L)\epsilon_t \quad (1)$$

où ϵ_t est un bruit blanc faible, L est l'opérateur retard, et ϕ_k et θ_k sont des polynômes de degrés respectifs p et q :

$$\phi(L) = 1 - \phi_1 L - \dots - \phi_p L^p$$

$$\theta(L) = 1 + \theta_1 L + \dots + \theta_q L^q$$

L'équation (1) est l'unique solution stationnaire causale à condition que les racines de $\phi(x) = 0$ soient en dehors du cercle unitaire. Le processus est inversible à condition $\theta(x) = 0$ soient en dehors du cercle unitaire.

1.10.4 Le processus intégrée I(d)

Une série X_t est intégrée d'ordre d , noté $I(d)$, si

$$\nabla^d x_t = \epsilon_t$$

Où $\nabla^d \equiv (1 - L)^d$ où L est l'opérateur retard. Une série X_t est intégrée d'ordre d si :

$$(1 - L)^d x_t = \epsilon_t$$

1.10.5 Le processus ARIMA(p,d,q)

La condition de stationnarité des modèles présentés ici n'est évidemment pas toujours convenable. On peut intégrer certains types de non-stationnarités en élargissant le modèle ARMA. Ces élargissements consistent en une série d'opérations préalables visant à éliminer la tendance ou la saisonnalité.

Définition 1.9 (ARIMA(p,d,q)). *Le processus (X_t) est un processus ARIMA $_{(p,d,q)}$ si le processus*

$$Y_t = \nabla^d X_t = (1 - L)^d X_t.$$

et un processus Y_t est un processus ARMA $_{(p,q)}$.

Le modèle ARIMA revient à appliquer un modèle ARMA sur le processus différencié :

$$Y_t = ARMA_{(p,q)} \iff X_t = ARIMA_{(p,d,q)}$$

L'équation d'un modèle $ARIMA_{(p,d,q)}$ est donc donnée par :

$$\phi(L)Y_t = \theta(L)\epsilon_t \iff \phi(L)\nabla^d X_t = \theta(L)\epsilon_t$$

Où ϕ et θ sont deux polynômes de degrés respectifs p et q .

Le "I" de $ARIMA$ signifie "integrated" comme réciproque de la différenciation.

Evidemment, le degré d n'est généralement pas connu. Pour le déterminer on peut agir par tâtonnements ou avoir recours à des tests de stationnarité : puisqu'un processus $ARMA_{(p,q)}$ est stationnaire, on cherche d tel qu'on puisse accepter l'hypothèse de stationnarité pour le processus $Y_t = \nabla^d X_t$. De façon générale, on se réfère à un principe de parcimonie et cherche la valeur satisfaisante minimale de d .

Il y a aussi d'autres modèles de séries chronologiques intéressantes parmi : les SARIMA, ARMAX, ARCH, GARCH,...

Chapitre 2

Les modèles de rupture

2.1 Introduction

La fréquence des changements de situation dans plusieurs domaines scientifiques explique l'intérêt porté à l'analyse statistique des points de rupture et de l'estimation. En effet, dans la pratique, lorsque les structures de contrôle révèlent qu'il y'a des ruptures quelque part dans l'évolution du phénomène étudié, il est naturel que l'on veuille localiser la position de ces ruptures qui est à l'origine du changement de régime. un lieu ou un temps de sorte que les observations suivent une distribution jusqu'à ce point et suivent une autre distribution après celui-ci. Plusieurs problèmes peuvent être définis similairement. Ainsi l'approche est double : on peut se contenter de vérifier s'il y'a rupture (souvent considéré comme un problème de test d'hypothèse) ou bien de localiser le point de rupture s'il y'a lieu (vu parfois comme un problème d'estimation).

Une rupture d'un modèle paramétrique est définie par un changement de l'une ou plusieurs de ces caractéristiques. Il peut s'agir par exemple de la moyenne, de la variance, du moment d'ordre r , de la famille de la distribution...

Quand ce modèle paramétrique change de structure, il est important de connaître le point où a eu lieu le changement ainsi que les valeurs des paramètres du modèle avant et après la rupture.

Définition 2.1. Point de rupture

Nous dirons qu'il y'a rupture dans une suite de variables aléatoires $X_1, X_2, X_3, \dots, X_n$ s'il existe un entier m ($1 \leq m \leq n - 1$) ; tel que les variables aléatoires $X_1, X_2, X_3, \dots, X_m$ suivent une loi de probabilité de fonction de répartition $F_1(x | \theta_1)$ et les variables aléatoires X_{m+1}, \dots, X_n suivent une autre loi de probabilité de fonction de répartition $F_2(x | \theta_2)$, ou θ_1 et θ_2 sont des paramètres inconnus réels ou vectoriels ($\theta_1 \neq \theta_2$) .

Le point m est appelé point de rupture ou point de changement . Dans la littérature anglo-saxonne il est connu sous le nom de shift point ou change point.

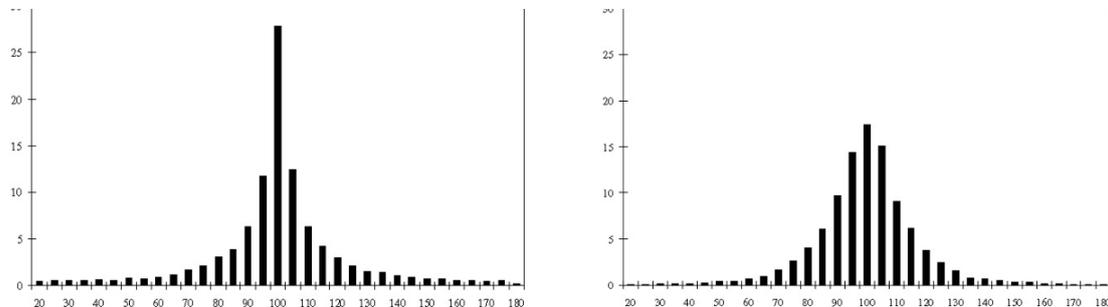
2.2 Les différents modèles de rupture

2.2.1 Rupture unique

Nous dirons qu'il y a rupture dans une suite de variables aléatoires X_1, \dots, X_n de fonction de répartition $F_i(x | \theta_i)$, $i = \overline{1, n}$, s'il existe un entier m ($1 \leq m \leq n - 1$) telque :

$$X_i = \begin{cases} F_1(x | \theta_1), & i = 1, \dots, m \\ F_2(x | \theta_2), & i = m + 1, \dots, n \end{cases}$$

Avec θ_1 et θ_2 sont des paramètres inconnus réels ou vectoriels ($\theta_1 \neq \theta_2$) .

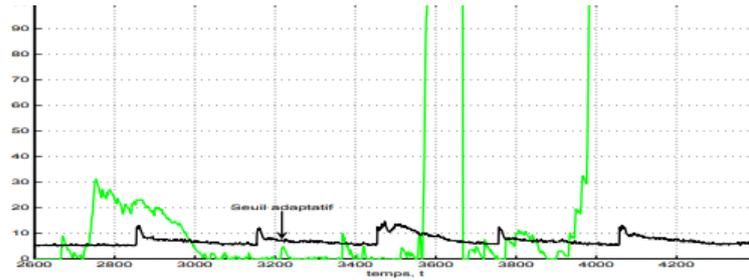


la figure représente une rupture dans une loi normale.

2.2.2 Ruptures multiples

Nous dirons qu'une suite de variables aléatoires admet des ruptures multiples s'il existe des entiers m, \dots, q , avec $(1 \leq m \leq \dots \leq q \leq n - 1)$ telque :

$$\theta_1 = \dots = \theta_m \neq \theta_{m+1} = \dots = \theta_q \neq \theta_{q+1} = \dots = \theta_n$$



La figure représente une séries ayant subi plusieurs rupture par rapport au seuil adaptatif.

2.2.3 Rupture épidémique

Soit X_1, \dots, X_n des variables aléatoires réelles de repartition continues respective F_1, \dots, F_n . Considérons le modèle épidémique suivant :

Il existe des entiers m et q ($1 \leq m \leq q \leq n$) tels que

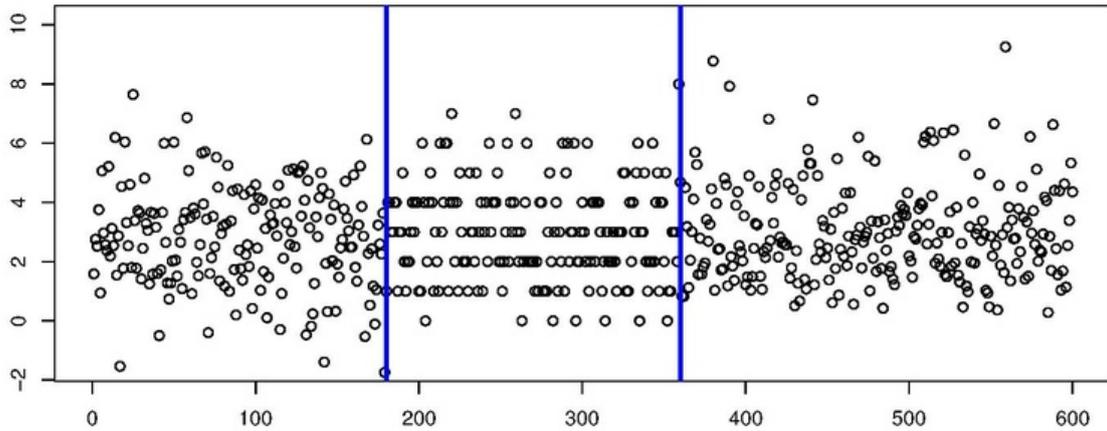
$$F_1 = F_2 = \dots = F_m = F_{q+1} = \dots = F_n = F$$

$$F_{m+1} = \dots = F_q = G \text{ et } F \neq G$$

Nous notons $l = q - m$ la longueur de l'épidémie.

Et m, q les bornes de localisation.

Parmi les chercheurs qui se sont intéressés au problèmes de rupture épidémique, citons Rackauskas et Suquet .



La figure indique une rupture épédémique.

2.3 Tests statistiques de rupture

Le plus souvent, les procédures considérées suppose que le point de rupture correspond à l'un des instants d'observations et tentent de répondre à la question suivante :

Y a-t-il une rupture dans le modèle considéré ou non ?

Ces procédures sont en général basées sur les tests d'hypothèses. dans le cas ou l'hypothèse nulle est rejetée, certains auteurs s'intéressent à l'estimation de ce point de rupture.

2.3.1 Tests de ruptures multiples

Considérons une suite de variables aléatoire X_1, \dots, X_n de fonction de repartition $F_i(X | \theta_i)$ de paramètres $\theta_i \quad i = 1, \dots, n$.

Le problème est formulé par un test entre l'hypothèse nulle H_0 , où il n'y a pas de ruptures, et l'hypothèse alternative H_1 , où des ruptures sont observées aux instants T_m, \dots, T_q :

$$H_0 : \quad \theta_1 = \theta_2 = \dots = \theta_n$$

$$H_1 : \theta_1 = \dots = \theta_{T_m} \neq \theta_{T_m+1} = \dots = \theta_{T_q} \neq \theta_{T_q+1} = \dots = \theta_n$$

Il existe plusieurs hypothèses de la forme H_1 , on considère pour commencer que le nombre et les positions des ruptures T_m, \dots, T_q sont connues. L'objectif est de construire une méthode qui permette de choisir entre H_0 et H_1 en fonction du vecteur X . Le test statistique compare l'adéquation des observations au modèle sous H_0 , où il n'y a pas de rupture, avec celle au modèle sous H_1 . Une règle de décision est établie à partir d'une statistique de test Z pour trancher entre les deux hypothèses, sous la forme :

$$\begin{cases} \text{Accepter } H_0; & \text{si } Z > \alpha \\ \text{Rejeter } H_0; & \text{si } Z \leq \alpha \end{cases}$$

Pour un seuil α approprié. En prenant une décision, le test risque de se tromper de deux manières possibles.

Définition 2.2. (*Erreur de type I*). *L'erreur de type I est faite lorsqu'on détecte à tort une rupture qui n'existe pas. Elle est également appelée risque de première espèce ou niveau, et est notée α*

Définition 2.3. (*Erreur de type II*). *L'erreur de type II est faite lorsqu'on ne détecte pas une vraie rupture. Elle est également appelée risque de deuxième espèce ou puissance du test, et est notée β .*

Pour garantir l'efficacité du test, il faut limiter ces risques. Ils sont quantifiés par les probabilités suivant :

$$\begin{aligned} P_1 &= P(H_0 \text{ est rejetée} \mid H_0 \text{ est vraie}) \\ &= \alpha \end{aligned}$$

$$\beta = P(\text{rejetée } H_1 \mid H_1 \text{ est vraie})$$

$$\begin{aligned} P_2 &= P(H_0 \text{ est rejetée} \mid H_0 \text{ est fausse}) \\ &= 1 - \beta \end{aligned}$$

qui est la puissance du test

Lorsque ces probabilités s'expriment en fonction d'un paramètre du test, de préférence ajustable, il est possible de déterminer les conditions pour lesquelles un nombre donné d'erreurs de type I et II sont observées en probabilité. On parle alors de contrôle de la détection. Plusieurs tests peuvent être construits pour un même problème. Pour choisir celui qui conduira à la meilleure détection des ruptures, il faut déterminer quel est le test le plus puissant au niveau α souhaité, c'est-à-dire celui qui maximise la probabilité de détection P_2 pour une probabilité P_1 donnée.

2.3.2 Tests de rupture unique à une position donnée

Le point X_{T_m} est l'unique rupture, où $X_i \in \mathbb{R}$, pour tout $1 \leq i \leq n$. Donc le problème devient :

$$H_0 : \theta_1 = \theta_2 = \dots = \theta_n \quad \text{contre} \quad H_1 : \theta_1 = \theta_2 = \dots = \theta_{T_m} \neq \theta_{T_m+1} = \dots = \theta_n. \quad (2.1)$$

Cette formulation est équivalente à celle d'un test d'homogénéité entre les deux échantillons $Y_1 = (X_1, \dots, X_{T_m})$ et $Y_2 = (X_{T_m+1}, \dots, X_n)$. Les tests classiques pour la détection d'un saut de moyenne, ou assimilé, comme les tests de Student, de Welch, d'Hotelling ou de Wilcoxon. Un grand nombre d'entre eux repose sur l'hypothèse que les données suivent une distribution normale. Quand il y'a rupture dans la moyenne.

Le problème de tester la stationarité (rupture contre pas de rupture) du modèle (2,1) à fait l'objet de plusieurs études utilisant soit l'approche classique de l'inférence fondée sur la fonction de vraisemblance, soit l'approche Bayésienne basée sur les distribution a posteriori , ou encore les méthodes non paramétriques dont le pionnier fut Page(1954,1955,1957). Ce dernier a développé des méthodes basées sur les sommes cumulatives, appelées techniques de Cusum, pour détecter un changement dans les

distributions d'une suite de variables aléatoire indépendantes.

Quandt (1958,1960) a utilisé le test de rapport de vraisemblance en conditionnant l'estimateur du maximum de vraisemblance du point de rupture.

Chernoff et Zacks (1964) et Kander et Zacks (1966) ont étudié une séquence de variables aléatoires gaussiennes et proposé un test Bayésien pour détecter un changement dans la moyenne.

Brown, Durbin et Evans (1975) ont proposé des tests basés sur les sommes cumulatives et cumulatives des carrés des résidus, appelés respectivement test de Cusum et Cusum des carrés, pour tester la stationnarité des séries.

Tsurumi (1977), Choy et Broemeling (1980) et Holbert (1982) ont proposé le traditionnel de Bayes, i.e. le rapport des probabilités a posteriori en affectant à l'hypothèse nulle $m = n$ (pas de rupture) une probabilité a priori constante et à l'hypothèse alternative une probabilité a priori uniforme sur $1, \dots, n$. Ce test est, en fait, sensible aux choix des lois a priori et pour ce, la robustesse est concernée.

Lindely (1985) et Box et Tiao(1965,1973) ont utilisé le test de signification Bayésien par la construction des régions de plus grande densité a posteriori (Highest posterior density, H.P.D), et souligné que ce type de test n'est approprié que lorsque l'information a priori sur les paramètres est vague et l'une des sous hypothèses simple.

2.3.3 Exemple de rupture dans la moyenne dans le cas des données gaussiennes

Nous nous intéressons à tester l'apparition d'un changement dans les moyennes d'une suite de variables aléatoires X_1, \dots, X_n gaussiennes et indépendantes de variance constante

connue ou inconnue σ^2 .

Les hypothèses de test peuvent s'écrire comme suit :

$$H_0 : \mu_1 = \mu_2 = \dots = \mu_n$$

$$H_1 : \mu_1 = \mu_2 = \dots = \mu_m \neq \mu_{m+1} = \mu_{m+2} = \dots = \mu_n$$

Avec m le point de rupture $1 \leq m < n$.

Pusieurs statistiques ont été proposées pour ce test, et parmi ces statistiques on retrouve celle du test du rapport de vraisemblance, celles suggérées par Chernoff et Zacks en (1964), Brown et Al 1975 et Pettitt en 1980.

Statistique du rapport de vraisemblance

Posons : $S_i = x_1 + x_2 + \dots + x_i$ pour tout $i = 1, 2, \dots, n$.

La racine carrée du log de la statistique du rapport de vraisemblance pour ce test est donnée par :

$$\max_{1 \leq K < n} [|KS_n/n - S_K| / K(1 - K/n)^{\frac{1}{2}}]$$

L'expression $[|KS_n/n - S_K| / K(1 - K/n)^{\frac{1}{2}}]$ représente la différence entre les moyennes des K premières observations et les $n - K$ dernières, c'est à dire d'égalité de deux moyennes de deux échantillons gaussiens.

Le maximum est recherché pour retrouver le point qui sépare l'échantillon en deux sous-échantillons de moyennes différentes.

Ce test est invariant si on change l'emplacement des observations, on peut aussi réduire le nombre d'observations et dans ce cas le test ne dépend pas directement des x_i mais des différences $y_i = x_i - x_1$, $i = 1, 2, \dots, n$.

Pour des valeurs données de m et de $\delta = \mu_i - \mu_1$, le rapport de vraisemblance de y sous H_0 et H_1 est :

$$\exp[\delta(mS_n/n - S_m) - 1/2m(1 - m/n)\delta^2]$$

En s'intéressant au test unilatéral, on suppose que le signe de δ est connu, voir $\delta > 0$, dans ce cas la statistique du rapport de vraisemblance peut s'écrire sans valeur absolue et donc la généralisation suivante peut être considérée comme une statistique de ce test :

$$\max_{n_0 \leq K < n_1} [KS_n/n - S_K/K(1 - K/n)^{\frac{1}{2}}]$$

Avec $1 \leq n_0 < n_1 < n$.

Statistique de Pettitt, 1980

Reprenons la statistique précédente :

$$\exp[\delta(mS_n/n - S_m) - 1/2m(1 - m/n)\delta^2]$$

En dérivant le log de cette dernière expression, avec $\delta = 0$, on obtient la statistique suggérée par Pettitt en 1980 qui est :

$$\max_{1 \leq K \leq n} (KS_n/n - S_K)$$

Statistique de Brown et Al, 1975

Une autre statistique intéressante pour ce test est celle proposée par Brown et Al en 1975 qui est défini comme suit :

$$\max_{1 < K \leq n} (\tilde{S}_{n-1} - \tilde{S}_{n-K}) / (K - 1)^{\frac{1}{2}}$$

Avec :

$$\tilde{S}_i = z_1 + z_2 + \dots + z_i$$

où : $z_i = i/(i + 1)^{1/2}(x_{i+1} - \bar{x}_i)$, on remarque que sous H_0 , les z_n sont des variables aléatoires indépendantes de loi normale centrée et réduite.

Statistique de Chernoff et Zacks, 1964

En supposant que m possède une loi a priori qui est la loi uniforme sur $1, 2, \dots, n$ et que δ tend vers zero, Chernoff et Zacks en 1964 ont proposé une statistique pour ce test dite "statistique quasi-bayésienne" qui est :

$$\sum_{i=1}^{n-1} i(i+1)^{1/2} \cdot z_i$$

2.4 Approche Bayésienne dans les modèles de rupture

Plusieurs approches sont utilisées dans les problèmes de rupture. Soit l'approche classique de l'inférence fondée sur la fonction de vraisemblance, où encore les méthodes non paramétriques basées sur les sommes cumulatives. Mais l'approche Bayésienne a été largement sollicitée. Cette méthodologie consiste à assigner des lois a priori (propres ou non informatives) aux paramètres inconnus du modèle, afin de déterminer les lois de probabilité a posteriori des paramètres d'intérêt et cela en utilisant le théorème de Bayes qui stipule que la loi de probabilité a posteriori d'un paramètre θ est proportionnelle au produit de la fonction de vraisemblance et de sa loi a priori :

$$\pi(\theta/S_n) \propto L(\theta/S_n) \cdot \pi(\theta)$$

où : θ est un paramètre vectoriel.

$S_n = (X_1, X_2, \dots, X_n)$ échantillon des observations.

$\pi(\theta)$ densité a priori du paramètre θ .

$\pi(\theta/S_n)$ densité a posteriori.

$L(\theta | S_n)$ la fonction de vraisemblance.

Par des intégrations successives, nous obtenons les différentes densités a posteriori marginales et conditionnelles des différents paramètres d'intérêt, qui nous permettent de déterminer l'estimateur Bayésien en utilisant différentes fonctions coût. (qui sera détaillé plus loin)

Dans cette section, nous citons quelques résultats d'estimation d'un point de rupture de certain modèles par la méthodologie Bayésienne. Dans les exemples suivants, les auteurs utilisent la fonction coût 0-1, ils estiment que le mode de la densité a posteriori marginale $\pi(m/S_n)$ détecte mieux le point de rupture. (voir chapitre 3)

2.4.1 Modèle de rupture dans le cas où les variables sont indépendantes

Présentation du modèle

Considérons le modèle suivant :

$$\begin{cases} X_i = \phi_0 + \epsilon_i; & i = 1, 2, \dots, m \\ X_i = \phi_1 + \epsilon_i; & i = m + 1, \dots, n \end{cases} \quad (2.2)$$

Où :

ϕ_0 et ϕ_1 sont des constantes réelles, inconnues et différentes. Elles représentent la moyenne des variables aléatoires X_i avant et après le changement respectivement.

Les ϵ_i représentent les erreurs aléatoires, elles sont gaussiennes, indépendantes et identiquement distribuées de moyenne nulle et de variance constante et inconnue σ^2 ; ($\epsilon_i \sim N(0, \sigma^2)$).

Le paramètre m représente le point de rupture. C'est un entier qui varie entre 1 et $n - 1$, n étant la taille de l'échantillon.

En assignant des lois a priori impropres (ou non informatives) aux paramètres σ^2 et (ϕ_0, ϕ_1) et une loi uniforme sur $1, 2, \dots, n - 1$ pour le point de rupture m , le théorème suivant nous détermine la masse de probabilité marginale a posteriori du point de rupture m .

Théorème 2.1. (L.D.Broemeling et D.Holbert, 1977)

Etant donné le modèle (2.2)

$$\begin{cases} X_i = \phi_0 + \epsilon_i; & i = 1, 2, \dots, m \\ X_i = \phi_1 + \epsilon_i; & i = m + 1, \dots, n \end{cases}$$

Où : $\phi_0 \neq \phi_1$ et $\epsilon_i \sim N(0, \sigma^2)$; ($\sigma > 0$) avec ϕ_0, ϕ_1 et σ sont des paramètres inconnus.

Si les densités a priori des paramètres $\sigma^2, (\phi_0, \phi_1)$ et du point de rupture m sont données respectivement par :

$$\pi_0(\sigma^2) \propto 1/\sigma^2$$

$$\pi_0(\phi_0, \phi_1) \propto \text{constante sur } \mathbb{R}^2$$

$$\pi_0(m) \propto 1/(n-1) \text{ pour } m = 1, 2, \dots, n-1$$

et $\sigma^2, (\phi_0, \phi_1)$ et m sont indépendants.

Alors la masse de probabilité a posteriori du point de rupture m est donnée par :

$$\pi_1(m) = \pi(m/X) \propto [m(n-m)]^{-1/2} [S_1^m + S_{m+1}^n]^{-(n-2)/2}$$

$$\text{où : } S_j^k = \sum_{i=j}^k (X_i - \bar{X}_j^k)^2$$

$$\text{et } \bar{X}_j^k = \frac{1}{k-j+1} \sum_{i=j}^k X_i, \text{ pour } k = 1, \dots, n.$$

2.4.2 Modèle de rupture dans le cas où les variables sont dépendantes

Dans cette partie, nous considérons quelques modèles de rupture dont les observations sont dépendantes. Dans un premier temps, nous étudions le modèle précédent dans le cas

où les erreurs sont autocorrélées, ensuite nous nous intéressons aux modèles de régression linéaire ou d'autorégression de l'erreur ayant subi un changement dans les paramètres de régression.

2.4.2.1 Modèle avec des erreurs autocorrélées

Reprenons le modèle (2.2) et supposons que les moyennes ϕ_0 et ϕ_1 sont différentes et inconnues.

Les $\epsilon_i; i = 1, \dots, n$ sont régies par un processus autorégressif d'ordre un (AR1).

$\epsilon_i = \rho\epsilon_{i-1} + e_i; e_i \sim N(0, \sigma_i^2); i = 1, 2, \dots, n; \rho$ et σ_i^2 sont inconnus.

La dépendance des variables $X_i; i = 1, \dots, n$ intervient dans la corrélation des erreurs $\epsilon_i; i = 1, \dots, n$

En assignant une densité conjointe a priori impropre aux paramètres m, ϕ_0, ϕ_1, ρ et σ^2 , le théorème suivant nous détermine la masse de probabilité marginale a posteriori du point de rupture m .

Théorème 2.2. (L.D.Broemeling et D.Holbert,1977)

Considérons le modèle (2.1) :

$$\begin{cases} X_i = \phi_0 + \epsilon_i; & i = 1, 2, \dots, m \\ X_i = \phi_1 + \epsilon_i; & i = m + 1, \dots, n \end{cases}$$

où :

$\epsilon_i = \rho\epsilon_{i-1} + e_i; e_i \sim N(0, \sigma^2); i = 1, 2, \dots, n.$

ρ et σ^2 sont inconnus.

Si on suppose que la densité conjointe a priori des paramètres m, ϕ_0, ϕ_1, ρ et σ^2 est :

$$\pi_0(m, \phi_0, \phi_1, \rho, \sigma^2) \propto 1/((n-1)\sigma^2)$$

alors la masse de probabilité a posteriori du point de rupture m est donnée par :

$$\pi_1(m) = \int_{\mathbb{R}} \frac{1}{\sqrt{\alpha_1 a - \beta_1^2}} \cdot \left[\alpha_3 - \frac{\beta_2^2}{a} - \left((\alpha_2 a + \beta_1 \beta_2) / \sqrt{\alpha_1 a^2 - \beta_1^2 a} \right)^2 \right]^{-(n/2-1)} d\rho$$

où :

$$a = m(1 - \rho)^2 + \rho^2,$$

$$\alpha_1 = (n - m - 1)(1 - \rho)^2 + 1,$$

$$\alpha_2 = (1 - \rho) \sum_{i=m+2}^n (X_i - \rho X_{i-1}) + (X_{m+1} - \rho X_m),$$

$$\alpha_3 = \sum_{i=2}^n (X_i - \rho X_{i-1})^2,$$

$$\beta_1 = \rho,$$

$$\beta_2 = (1 - \rho) \sum_{i=2}^m (X_i - \rho X_{i-1}) + \rho(X_{m+1} - \rho X_m).$$

2.4.2.2 Modèle de régression linéaire

Généralement, les problèmes de rupture sont représentés par des modèles de régression. Dans cette section, nous considérons un modèle de régression linéaire avec un changement dans les paramètres de régression.

Présentation du modèle

Le modèle considéré est le suivant :

$$\begin{cases} Y_i = X_i\beta_1 + \epsilon_i, & i = 1, 2, \dots, m \\ Y_i = X_i\beta_2 + \epsilon_i, & i = m + 1, \dots, n \end{cases} \quad (2.3)$$

où :

X_i ; $i = 1, \dots, n$ est un vecteur à $(1 \times p)$ observations constitué de p variables indépendantes.

Y_i est la i ème réalisation de la variable dépendante.

β_1 et β_2 deux vecteurs $(p \times 1)$, ils représentent les paramètres de la régression, avec $(\beta_1 \neq \beta_2)$.

ϵ_i ; $i = 1, \dots, n$ sont les erreurs du modèle, elles sont supposées d'une part indépendantes, identiquement distribuées suivant une loi normale $N(0, \sigma^2)$, et d'autre part autocorrélées, plus précisément, elles sont régies par un processus autorégressif d'ordre p , $p \geq 1$, $(AR(p))$.

m ; point de rupture, c'est un paramètre aléatoire prenant les valeurs $1, 2, \dots, n - 1$. Le changement concerne uniquement les paramètres de régression.

Cas où les erreurs sont indépendantes

Ce modèle a été considéré par J.H. Chin choy et L.D. Broemeling en 1980. Le théorème suivant nous donne les densités a posteriori de m .

Théorème 2.3. (CHIN.CHOY et BROEMELING,1980)

Etant donné le modèle (2.3).

m, β, σ^2 sont des paramètres inconnus, avec $\beta = (\beta'_1, \beta'_2)$.
 m est uniformément distribué sur l'ensemble $\{1, 2, \dots, n - 1\}$

La loi conjointe a priori de β et de $R = 1/\sigma^2$ est :

$$\pi_0(B/R = r) \equiv 2p - N(\beta_\mu, r\tau)$$

tel que $\beta_\mu \in \mathbb{R}^{2p}$, et τ est une matrice $(2p \times 2p)$ donnée, symétrique, définie positive.

Une telle loi est appelée une loi normale-Gamma.

R suit une distribution Gamma de paramètres a, b ($a > 0, b > 0$).

m est indépendante de β et de R .

Alors la masse de probabilité a posteriori du point de rupture m est :

$$\pi(m/y) \propto \begin{cases} D(m)^{-a^*} |x'(m)x(m) + \tau|^{-1/2}; & m = 1, 2, \dots, n-1 \\ 0 & \text{sinon} \end{cases}$$

où $a^* = a + n/2$.

$$\begin{aligned} D(m)^{-a^*} &= b + 1/2 \{y'y + \beta'_\mu \tau \beta_\mu - \beta'^*(m) [x'(m)x(m) + \tau] \beta^*(m)\} \\ &= b + 1/2 \{[y - x(m)\beta^*(m)]' . y + [\beta_\mu - \beta^*(m)]' \tau \beta_\mu\} \end{aligned}$$

Cas où les erreurs sont autocorrélées

Considérons le même modèle de régression et intéressons nous au cas où les erreurs sont régies par un processus autorégressif d'ordre un ($AR(1)$).

Soit le théorème suivant :

Théorème 2.4. (SALAZAR et BROEMELING,1981) *Considérons le modèle suivant :*

$$\begin{cases} Y_i = X_i \beta_1 + \mu_i; & i = 1, 2, \dots, m \\ Y_i = X_i \beta_2 + \mu_i; & i = m + 1, \dots, n \end{cases}$$

$$\mu_i = \rho\mu_{i-1} + e_i; \quad i = 1, \dots, n$$

où $e_i \sim iid \ N(0, r); \quad 0 < r = \sigma^{-2}$

$m, \beta_1, \beta_2, \rho,$ et $r = \sigma^{-2}$ inconnus.

Y_i est la i ème observation de la variable dépendante.

X_i est la i ème observation de K variables indépendantes.

$\rho \in \mathbb{R}$, il est muni d'une distribution non informative sur \mathbb{R} ($\rho \propto cste$ sur \mathbb{R}).

m : point de rupture, uniformément distribuée sur $I_{n-1} = \{1, 2, \dots, n-2\}$

$\beta = (\beta'_1, \beta'_2) \in \mathbb{R}^{2K}$: coefficient de régression. ($\beta_1 \neq \beta_2$).

La distribution conjointe a priori pour $\beta_i (i = 1, 2)$ et R est :

$\pi_0(\beta/R = r) \equiv N(\mu, rP)$ c'est la loi normale $2K$ -dimensionnelle.

$\mu = (\mu'_1, \mu'_2)' \in \mathbb{R}^{2K}$, vecteur moyen.

P une matrice $2K \times 2K$ symétrique, définie positive.

La densité a priori de R est une Gamma de paramètres a et b strictement positifs.

$\pi_0(r) \propto r^{a-1} \exp(-br)$

On suppose de plus $m, (\beta, R), \rho$ indépendants.

Alors la masse de probabilité a posteriori du point de rupture m est donnée par :

$$\pi(m/Y) = \begin{cases} \int_{\mathbb{R}} |H(\rho)|^{-1/2} C(\rho)^{-(a+n/2)}; & m = 1, 2, \dots, n-2 \\ 0 & \text{ailleurs} \end{cases}$$

Où :

$$C(\rho) = 2b + \mu' P \mu + \sum_{t=1}^n (Y_t - \rho Y_t - 1)^2 \tilde{\beta}' H(\rho) \tilde{\beta}$$

$$Z_1 = ((X_1 - \rho X_0)', (X_2 - \rho X_1)', \dots, (X_m - \rho X_{m-1})')$$

$$Z_2 = ((X_{m+2} - \rho X_{m+1})', \dots, (X_n - \rho X_{n-1})')$$

$$V_1 = ((Y_1 - \rho Y_0)', (Y_2 - \rho Y_1)', \dots, (Y_m - \rho Y_{m-1})')$$

$$V_2 = ((Y_{m+2} - \rho Y_{m+1})', \dots, (Y_n - \rho Y_{n-1})')$$

$$\begin{aligned}
 H(\rho) &= X'X + P = \begin{pmatrix} H_{11} & H_{12} \\ H_{21} & H_{22} \end{pmatrix} \\
 &= \begin{pmatrix} P_{11} + Z_1'Z_1 + \rho^2 X_m'X_m & P_{12} - \rho X_m'X_m \\ P_{21} - \rho X_{m+1}'X_m & P_{22} + Z_2'Z_2 + X_{m+1}'X_m \end{pmatrix} \\
 P &= \begin{pmatrix} P_{11} & P_{12} \\ P_{21} & P_{22} \end{pmatrix}, \quad \tilde{\beta} = (\tilde{\beta}_1', \tilde{\beta}_2')' \\
 \tilde{\beta}_1 &= H_{11.2}^{-1}(\rho)(\alpha_1 - H_{12}(\rho)H_{22}^{-1}(\rho)\alpha_2) \\
 \tilde{\beta}_2 &= H_{22.1}^{-1}(\rho)(\alpha_2 - H_{21}(\rho)H_{11}^{-1}(\rho)\alpha_1) \\
 P\mu + X'y &= \begin{pmatrix} P_{11} & P_{12} \\ P_{21} & P_{22} \end{pmatrix} \begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix} + \begin{pmatrix} Z_1' & -\rho X_m' & 0 \\ 0 & X_{m+1}' & Z_2' \end{pmatrix} \\
 &= \begin{pmatrix} P_{11}\cdot\mu_1 + P_{12}\cdot\mu_2 + Z_1'V_1 - \rho X_m'(Y_{m+1} - \rho Y_m) \\ P_{21}\cdot\mu_1 + P_{22}\cdot\mu_2 + Z_2'V_2 + X_{m+1}'(Y_{m+1} - \rho Y_m) \end{pmatrix} = \begin{pmatrix} \alpha_1 \\ \alpha_2 \end{pmatrix}
 \end{aligned}$$

Chapitre 3

Estimation Bayésienne d'un point de rupture dans un processus autorégressif

3.1 Introduction

L'analyse des séries temporelles, ou séries chronologiques, se réfère à la branche de statistique où les observations sont régulièrement espacées dans le temps. Leur domaine d'application est très vaste et s'étend de l'astronomie à l'économie et finance en passant par la biologie, psychologie, géophysique ou la théorie du signal, météorologie...etc . Lorsqu'un changement brusque apparaît dans ces différents domaines (exemple du réchauffement climatique en météorologie) , nous nous intéressons à un modèle de rupture dans une série chronologique.

Dans ce chapitre basé sur l'article de **Mayuri Pandya, Krichnam Batt et Hardik Pandya Chetan Thakar [2012]**, nous nous intéressons à l'estimation d'un point de rupture dans un processus autoregressif d'ordre un $AR(1)$, en utilisant l'approche Bayésienne.

Pour cela nous cherchons la densité a posteriori des paramètres du modèle étudié, puis

nous déduisons l'estimateur de Bayes en utilisant des fonctions coûts usuelles.

3.2 Présentation du modèle

Nous nous intéressons à un processus autoregressif d'ordre un (AR(1)), ayant subi un changement de structure à un instant inconnu m . Le changement concerne à la fois le coefficient d'autocorrélation et la variance des erreurs.

Le modèle s'écrit donc comme suit :

$$X_i = \begin{cases} \beta_1 X_{i-1} + \epsilon_i & , \quad i = 1, 2, \dots, m \\ \beta_2 X_{i-1} + \epsilon_i & , \quad i = m + 1, \dots, n. \end{cases}$$

Avec β_1, β_2 sont les coefficients d'autocorrélations supposés inconnus avec : ($\beta_1 \neq \beta_2$). x_i est la i^{eme} observation de la variable dépendante, $i = \overline{1, n}$ n étant la taille de l'échantillon.

ϵ_i : sont les erreurs du modèle, elles sont supposées indépendantes, identiquement distribuées suivant une loi normale $N(0, \sigma_i^2)$.

$$\epsilon_i \sim N(0, \sigma_1^2) \quad \text{pour } i = 1, 2, \dots, m$$

$$\epsilon_i \sim N(0, \sigma_2^2) \quad \text{pour } i = m + 1, \dots, n$$

Avec σ_1^2 et σ_2^2 connues.

m est le point de rupture, c'est un paramètre aléatoire prenant les valeurs $1, 2, \dots, n - 1$.

3.3 Estimation Bayésienne

La méthode du maximum de vraisemblance ainsi que d'autres approches classiques sont basés uniquement sur les informations empiriques fournies par les données.

Cependant, quand il y'a une connaissance technique des paramètres de la distribution ; la méthode Bayésienne apparaît comme une approche attrayante devant les difficultés théoriques des approches classiques.

La procédure de Bayes est basée sur la densité a posteriori jointe des différents paramètres $g(\beta_1, \beta_2 \text{ et } m | X)$, qui est proportionnelle au produit de la fonction de vraisemblance et la densité jointe a priori :

$$g(\beta_1, \beta_2, m | X) = \frac{L(\beta_1, \beta_2, m | X)g(\beta_1, \beta_2, m)}{\sum_{m=1}^{n-1} \int \int L(\beta_1, \beta_2, m | X)g(\beta_1, \beta_2, m)d\beta_1 d\beta_2}$$

Avec : $L(\beta_1, \beta_2, m | X)$ la fonction de vraisemblance des différents paramètres.

et : $g(\beta_1, \beta_2, m)$ la densité jointe a priori de $m, \beta_1 \text{ et } \beta_2$

3.3.1 La fonction de vraisemblance :

Soit :

$$X_i = \begin{cases} \beta_1 x_{i-1} + \epsilon_i & , \quad i = 1, 2, \dots, m \\ \beta_2 x_{i-1} + \epsilon_i & , \quad i = m + 1, \dots, n. \end{cases}$$

d'où :

$$\epsilon_i = \begin{cases} x_i - \beta_1 x_{i-1} & , \quad i = 1, 2, \dots, m \\ x_i - \beta_2 x_{i-1} & , \quad i = m + 1, \dots, n. \end{cases}$$

La densité de vraisemblance est :

$$L(\beta_1, \beta_2, m | X) = \prod_{i=1}^n f_{X_i}(x)$$

Nous avons : $\epsilon_i \sim N(0, \sigma_1^2)$, pour $i = 1, \dots, m$.

$\epsilon_i \sim N(0, \sigma_2^2)$, pour $i = m + 1, \dots, n$.

La densité des ϵ_i :

$$f(\epsilon_i) = \begin{cases} \frac{1}{\sqrt{2\pi}\sigma_1} e^{-\frac{1}{2\sigma_1^2}\epsilon_i^2} & , \quad i = 1, 2, \dots, m \\ \frac{1}{\sqrt{2\pi}\sigma_2} e^{-\frac{1}{2\sigma_2^2}\epsilon_i^2} & , \quad i = m + 1, \dots, n. \end{cases}$$

d'où :

$$f(X_i) = \begin{cases} \frac{1}{\sqrt{2\pi}\sigma_1} e^{-\frac{1}{2\sigma_1^2}(x_i - \beta_1 x_{i-1})^2} & , \quad i = 1, 2, \dots, m \\ \frac{1}{\sqrt{2\pi}\sigma_2} e^{-\frac{1}{2\sigma_2^2}(x_i - \beta_2 x_{i-1})^2} & , \quad i = m + 1, \dots, n. \end{cases}$$

Comme les ϵ_i sont indépendantes

Alors la fonction de vraisemblance est donnée par :

$$L(\beta_1, \beta_2, m | X) = \frac{1}{\sqrt{2\pi}\sigma_1} e^{-\frac{1}{2\sigma_1^2}(x_i - \beta_1 x_{i-1})^2} \dots \frac{1}{\sqrt{2\pi}\sigma_1} e^{-\frac{1}{2\sigma_1^2}(x_i - \beta_1 x_{i-1})^2} \cdot \frac{1}{\sqrt{2\pi}\sigma_2} e^{-\frac{1}{2\sigma_2^2}(x_i - \beta_2 x_{i-1})^2} \dots \frac{1}{\sqrt{2\pi}\sigma_2} e^{-\frac{1}{2\sigma_2^2}(x_i - \beta_2 x_{i-1})^2}$$

Ce qui nous donne :

$$\begin{aligned} L(\beta_1, \beta_2, m | X) &= \left(\frac{1}{\sqrt{2\pi}\sigma_1} \right)^m \exp \left[\frac{-1}{2\sigma_1^2} \sum_{i=1}^m (x_i - \beta_1 x_{i-1})^2 \right] \left(\frac{1}{\sqrt{2\pi}\sigma_2} \right)^{n-m} \exp \left[\frac{-1}{2\sigma_2^2} \sum_{m+1}^n (x_i - \beta_2 x_{i-1})^2 \right] \\ &= (\sqrt{2\pi}\sigma_1)^{-m} (\sqrt{2\pi}\sigma_2)^{-(n-m)} \exp \left[\frac{-1}{2\sigma_1^2} \left(\sum_{i=1}^m x_i^2 - 2\beta_1 \sum_{i=1}^m x_i x_{i-1} + \beta_1^2 \sum_{i=1}^m x_{i-1}^2 \right) \right] \\ &\quad \exp \left[\frac{-1}{2\sigma_2^2} \left(\sum_{m+1}^n x_i^2 - 2\beta_2 \sum_{m+1}^n x_i x_{i-1} + \beta_2^2 \sum_{m+1}^n x_{i-1}^2 \right) \right] \\ &= (2\pi)^{-n/2} \sigma_1^{-m} \sigma_2^{-(n-m)} \exp \left[-\frac{1}{2} \beta_1^2 \frac{\sum_{i=1}^m x_{i-1}^2}{\sigma_1^2} + \beta_1 \frac{\sum_{i=1}^m x_i x_{i-1}}{\sigma_1^2} - \frac{1}{2} \frac{\sum_{i=1}^m x_i^2}{\sigma_1^2} \right] \\ &\quad \exp \left[-\frac{1}{2} \beta_2^2 \frac{\sum_{i=1}^m x_{i-1}^2}{\sigma_2^2} + \beta_2 \frac{\sum_{i=1}^m x_i x_{i-1}}{\sigma_1^2} - \frac{1}{2} \frac{\sum_{i=1}^m x_i^2}{\sigma_2^2} \right] \end{aligned}$$

Posons :

$$\begin{aligned} S_{K_1} &= \sum_{i=1}^k x_{i-1}^2 \\ S_{K_2} &= \sum_{i=1}^k x_i x_{i-1} \\ S_{K_3} &= \sum_{i=1}^k x_i^2 \\ K_1 &= (2\pi)^{\frac{-n}{2}} \end{aligned}$$

Alors :

$$L(\beta_1, \beta_2, m | X) = K_1 \exp \left[-\frac{1}{2} \beta_1^2 \left(\frac{S_{m1}}{\sigma_1^2} \right) + \beta_1 \left(\frac{S_{m2}}{\sigma_1^2} \right) - \frac{S_{m3}}{2\sigma_1^2} \right] \sigma_1^{-m} \sigma_2^{-(n-m)} \\ \exp \left[-\frac{1}{2} \beta_2^2 \left(\frac{S_{n1} - S_{m1}}{\sigma_2^2} \right) + \beta_2 \left(\frac{S_{n2} - S_{m2}}{\sigma_2^2} \right) - \left(\frac{S_{n3} - S_{m3}}{2\sigma_2^2} \right) \right]$$

3.3.2 Les lois a priori

Les distributions a priori des paramètres inconnus sont attribuées comme suit :

Une loi normale de moyenne μ_i et de variance a_i , pour les β_i avec $i = 1, 2$.

Une loi uniforme discrète sur l'intervalle $[1, n - 1]$

Dans la littérature, c'est pratiquement la seule utilisée car elle est naturelle.

De plus nous supposons que les paramètres β_1, β_2 et m sont indépendants.

Ce qui nous donne :

$$\begin{cases} g_1(\beta_1) = \frac{1}{\sqrt{2\pi a_1}} e^{-\frac{1}{2a_1}(\beta_1 - \mu_1)^2}; a_1, \mu_1 > 0 \text{ et } \beta_1 \in \mathbb{R} \\ g_2(\beta_2) = \frac{1}{\sqrt{2\pi a_2}} e^{-\frac{1}{2a_2}(\beta_2 - \mu_2)^2}; a_2, \mu_2 > 0 \text{ et } \beta_2 \in \mathbb{R} \\ g_3(m) = \frac{1}{n-1} \end{cases}$$

la densité a priori jointe

La densité a priori jointe des paramètres β_1, β_2 et m est donnée par :

$$g(\beta_1, \beta_2, m) = g(\beta_1) * g(\beta_2) * g(m) \\ = \frac{1}{\sqrt{2\pi a_1}} e^{-\frac{1}{2a_1}(\beta_1 - \mu_1)^2} \frac{1}{\sqrt{2\pi a_2}} e^{-\frac{1}{2a_2}(\beta_2 - \mu_2)^2} \frac{1}{n-1} \\ = (2\pi \sqrt{a_1} \sqrt{a_2} (n-1))^{-1} e^{-\frac{1}{2a_1}(\beta_1^2 - 2\beta_1 \mu_1 + \mu_1^2)} e^{-\frac{1}{2a_2}(\beta_2^2 - 2\beta_2 \mu_2 + \mu_2^2)}$$

posons :

$$K_2 = (2\pi(n-1)\sqrt{a_1}\sqrt{a_2})^{-1}$$

donc :

$$g(\beta_1, \beta_2, m) = K_2 \exp \left[\frac{-1}{2a_1} (\beta_1^2 - 2\beta_1\mu_1 + \mu_1^2) \right] \exp \left[\frac{-1}{2a_2} (\beta_2^2 - 2\beta_2\mu_2 + \mu_2^2) \right]$$

3.3.3 la densité a posteriori jointe

En vertu du théorème de Bayes nous avons :

$$g(\beta_1, \beta_2, m | X) = \frac{L(\beta_1, \beta_2, m | X)g(\beta_1, \beta_2, m)}{\sum_{m=1}^{n-1} \int_{\beta_1} \int_{\beta_2} L(\beta_1, \beta_2, m | X)g(\beta_1, \beta_2, m)d\beta_1 d\beta_2}$$

posons :

$$h(X) = \sum_{m=1}^{n-1} \int_{\beta_1} \int_{\beta_2} L(\beta_1, \beta_2, m | X)g(\beta_1, \beta_2, m)d\beta_1 d\beta_2$$

donc :

$$\begin{aligned} g(\beta_1, \beta_2, m | X) &= h^{-1}(X)K_1K_2 \exp \left[-\frac{1}{2}\beta_1^2 \left(\frac{S_{m1}}{\sigma_1^2} \right) + \beta_1 \left(\frac{S_{m2}}{\sigma_1^2} \right) - \frac{S_{m3}}{2\sigma_1^2} \right] \sigma_1^{-m} \sigma_2^{-(n-m)} \\ &\quad \exp \left[-\frac{1}{2}\beta_2^2 \left(\frac{S_{n1} - S_{m1}}{\sigma_2^2} \right) + \beta_2 \left(\frac{S_{n2} - S_{m2}}{\sigma_2^2} \right) - \left(\frac{S_{n3} - S_{m3}}{2\sigma_2^2} \right) \right] \\ &\quad \exp \left[\frac{-1}{2}\beta_1^2 \left(\frac{1}{a_1} \right) + \beta_1 \left(\frac{\mu_1}{a_1} \right) - \frac{1}{2} \left(\frac{\mu_1^2}{a_1} \right) \right] \exp \left[\frac{-1}{2}\beta_2^2 \left(\frac{1}{a_2} \right) + \beta_2 \left(\frac{\mu_2}{a_2} \right) - \frac{1}{2} \left(\frac{\mu_2^2}{a_2} \right) \right] \\ &= \exp \left[-\frac{1}{2}\beta_1^2 \left(\frac{S_{m1}}{\sigma_1^2} + \frac{1}{a_1} \right) + \beta_1 \left(\frac{S_{m2}}{\sigma_1^2} + \frac{\mu_1}{a_1} \right) \right] \\ &\quad \exp \left[-\frac{1}{2}\beta_2^2 \left(\frac{S_{n1} - S_{m1}}{\sigma_2^2} + \frac{1}{a_2} \right) + \beta_2 \left(\frac{S_{n2} - S_{m2}}{\sigma_2^2} + \frac{\mu_2}{a_2} \right) \right] \exp \left[-\frac{1}{2} \left(\frac{\mu_1^2}{a_1} + \frac{\mu_2^2}{a_2} \right) \right] \\ &\quad \exp \left[-\frac{1}{2} (S_{m3}\sigma_1^{-2} + S_{n3}\sigma_2^{-2} - S_{m3}\sigma_2^{-2}) \right] K_1K_2 h^{-1}(X) \sigma_1^{-m} \sigma_2^{-(n-m)} \\ &= K_1K_2 \exp \left[-\frac{1}{2} \left(\frac{\mu_1^2}{a_1} + \frac{\mu_2^2}{a_2} \right) \right] \exp \left[-\frac{1}{2} S_{n3}\sigma_2^{-2} \right] h^{-1}(X) \\ &\quad \exp \left[-\frac{1}{2}\beta_1^2 \left(\frac{S_{m1}}{\sigma_1^2} + \frac{1}{a_1} \right) + \beta_1 \left(\frac{S_{m2}}{\sigma_1^2} + \frac{\mu_1}{a_1} \right) \right] \\ &\quad \exp \left[-\frac{1}{2}\beta_2^2 \left(\frac{S_{n1} - S_{m1}}{\sigma_2^2} + \frac{1}{a_2} \right) + \beta_2 \left(\frac{S_{n2} - S_{m2}}{\sigma_2^2} + \frac{\mu_2}{a_2} \right) \right] \\ &\quad \exp \left[-\frac{1}{2} S_{m3} (\sigma_1^{-2} - \sigma_2^{-2}) \right] \sigma_1^{-m} \sigma_2^{-(n-m)}. \end{aligned}$$

Posons :

$$\begin{aligned} A_{1m} &= \frac{S_{m1}}{\sigma_1^2} + \frac{1}{a_1} , \quad A_{1m} > 0 \\ B_{1m} &= \frac{S_{m2}}{\sigma_1^2} + \frac{\mu_1}{a_1} \\ A_{2m} &= \frac{S_{n1} - S_{m1}}{\sigma_2^2} + \frac{1}{a_2} , \quad A_{2m} > 0 \\ B_{2m} &= \frac{S_{n2} - S_{m2}}{\sigma_2^2} + \frac{\mu_2}{a_2} \\ K_3 &= K_1 K_2 \exp\left[-\frac{1}{2}\left(\frac{\mu_1^2}{a_1} + \frac{\mu_2^2}{a_2}\right)\right] \exp\left[-\frac{1}{2}S_{n3}\sigma_2^{-2}\right] \end{aligned}$$

Donc :

$$\begin{aligned} g(\beta_1, \beta_2, m | X) &= K_3 h^{-1}(X) \exp\left[-\frac{1}{2}\beta_1^2 A_{1m} + \beta_1 B_{1m}\right] \exp\left[-\frac{1}{2}\beta_2^2 A_{2m} + \beta_2 B_{2m}\right] \\ &\quad \exp\left[-\frac{1}{2}S_{3m}(\sigma_1^{-2} - \sigma_2^{-2})\right] \sigma_1^{-m} \sigma_2^{-(n-m)} \end{aligned}$$

Calcul de la marginale de $h(\mathbf{X})$:

$$\begin{aligned} h(X) &= \sum_{m=1}^{n-1} \int_{\beta_1} \int_{\beta_2} L(\beta_1, \beta_2, m | X) g(\beta_1, \beta_2, m) d\beta_1 d\beta_2 \\ h(X) &= K_3 \int_{-\infty}^{+\infty} \exp\left[-\frac{1}{2}\beta_1^2 A_{1m} + \beta_1 B_{1m}\right] \int_{-\infty}^{+\infty} \exp\left[-\frac{1}{2}\beta_2^2 A_{2m} + \beta_2 B_{2m}\right] \partial\beta_1 \partial\beta_2 \\ &\quad \exp\left[-\frac{1}{2}S_{3m}(\sigma_1^{-2} - \sigma_2^{-2})\right] \sigma_1^{-m} \sigma_2^{-(n-m)}. \end{aligned}$$

Calcul de l'intégrales I_1 :

$$I_1 = \int_{-\infty}^{+\infty} \exp\left[-\frac{1}{2}\beta_1^2 A_{1m} + \beta_1 B_{1m}\right] d\beta_1$$

Nous avons :

$$\begin{aligned} \exp\left[-\frac{1}{2}A_{1m} \left(\beta_1 - \frac{B_{1m}}{A_{1m}}\right)^2\right] &= \exp\left[-\frac{1}{2}A_{1m} \left(\beta_1^2 - 2\beta_1 \frac{B_{1m}}{A_{1m}} + \frac{B_{1m}^2}{A_{1m}^2}\right)\right] \\ &= \exp\left[-\frac{1}{2}\beta_1^2 A_{1m} + \beta_1 B_{1m} - \frac{B_{1m}^2}{2A_{1m}}\right] \\ &= \exp\left[-\frac{1}{2}\beta_1^2 A_{1m} + \beta_1 B_{1m}\right] \exp\left[-\frac{B_{1m}^2}{2A_{1m}}\right] \end{aligned}$$

d'où :

$$\begin{aligned}
 I_1 &= \int_{-\infty}^{+\infty} \exp \left[-\frac{1}{2} A_{1m} \left(\beta_1 - \frac{B_{1m}}{A_{1m}} \right)^2 \right] \exp \left[\frac{B_{1m}^2}{2A_{1m}} \right] d\beta_1 \\
 &= \exp \left[\frac{B_{1m}^2}{2A_{1m}} \right] \int_{-\infty}^{+\infty} \exp \left[-\frac{1}{2A_{1m}^{-1}} \left(\beta_1 - \frac{B_{1m}}{A_{1m}} \right)^2 \right] d\beta_1 \\
 &= \exp \left[\frac{B_{1m}^2}{2A_{1m}} \right] \frac{\sqrt{2\pi}}{\sqrt{A_{1m}}} \int_{-\infty}^{+\infty} \frac{1}{\sqrt{2\pi A_{1m}}} \exp \left[-\frac{1}{2A_{1m}^{-1}} \left(\beta_1 - \frac{B_{1m}}{A_{1m}} \right)^2 \right] d\beta_1 \\
 I_1 &= \frac{\sqrt{2\pi} e^{\frac{B_{1m}^2}{2A_{1m}}}}{\sqrt{A_{1m}}}
 \end{aligned}$$

Calcul de l'intégrale I_2

$$I_2 = \int_{-\infty}^{+\infty} \exp \left[-\frac{1}{2} \beta_2^2 A_{2m} + \beta_2 B_{2m} \right] d\beta_2$$

nous avons :

$$\begin{aligned}
 \exp \left[-\frac{1}{2} A_{2m} \left(\beta_2 - \frac{B_{2m}}{A_{2m}} \right)^2 \right] &= \exp \left[-\frac{1}{2} A_{2m} \left(\beta_2^2 - 2\beta_2 \frac{B_{2m}}{A_{2m}} + \frac{B_{2m}^2}{A_{2m}^2} \right) \right] \\
 &= \exp \left[-\frac{1}{2} \beta_2^2 A_{2m} + \beta_2 B_{2m} - \frac{B_{2m}^2}{2A_{2m}} \right] \\
 &= \exp \left[-\frac{1}{2} \beta_2^2 A_{2m} + \beta_2 B_{2m} \right] \exp \left[-\frac{B_{2m}^2}{2A_{2m}} \right]
 \end{aligned}$$

d'où :

$$\begin{aligned}
 I_2 &= \int_{-\infty}^{+\infty} \exp \left[-\frac{1}{2} A_{2m} \left(\beta_2 - \frac{B_{2m}}{A_{2m}} \right)^2 \right] \exp \left[\frac{B_{2m}^2}{2A_{2m}} \right] d\beta_2 \\
 &= \exp \left[\frac{B_{2m}^2}{2A_{2m}} \right] \int_{-\infty}^{+\infty} \exp \left[-\frac{1}{2A_{2m}^{-1}} \left(\beta_2 - \frac{B_{2m}}{A_{2m}} \right)^2 \right] d\beta_2 \\
 &= \exp \left[\frac{B_{2m}^2}{2A_{2m}} \right] \frac{\sqrt{2\pi}}{\sqrt{A_{2m}}} \int_{-\infty}^{+\infty} \frac{1}{\sqrt{2\pi A_{2m}}} \exp \left[-\frac{1}{2A_{2m}^{-1}} \left(\beta_2 - \frac{B_{2m}}{A_{2m}} \right)^2 \right] d\beta_2 \\
 I_2 &= \frac{\sqrt{2\pi} e^{\frac{B_{2m}^2}{2A_{2m}}}}{\sqrt{A_{2m}}}
 \end{aligned}$$

d'où :

$$h(X) = K_3 \sum_{m=1}^{n-1} T_1(m)$$

avec :

$$T_1(m) = I_1(m)I_2(m)\exp\left[-\frac{1}{2}S_{m3}(\sigma_1^{-2} - \sigma_2^{-2})\right]\sigma_1^{-m}\sigma_2^{-(n-m)}.$$

$$\text{et : } K_3 = K_1K_2\exp\left[-\frac{1}{2}\left(\frac{\mu_1^2}{a_1} + \frac{\mu_2^2}{a_2}\right)\right]\exp\left[-\frac{1}{2}S_{n3}\sigma_2^{-2}\right]$$

Donc :

$$g(\beta_1, \beta_2, m | X) = [\exp(-\frac{1}{2}\beta_1^2 A_{1m} + \beta_1 B_{1m})\exp(-\frac{1}{2}\beta_2^2 A_{2m} + \beta_2 B_{2m})\exp(-\frac{1}{2}S_{3m}(\sigma_1^{-2} - \sigma_2^{-2}))\sigma_1^{-m}\sigma_2^{-(n-m)}] \left[\sum_{m=1}^{n-1} T_1(m)\right]^{-1}$$

3.3.4 Les densités a posteriori marginales

La distribution a posteriori marginale du point β_1

$$\begin{aligned} g(\beta_1 | X) &= \sum_{m=1}^{n-1} \int_{-\infty}^{+\infty} g(\beta_1, \beta_2, m | X) d\beta_2 \\ &= K_3 \sum_{m=1}^{n-1} \exp\left[-\frac{1}{2}\beta_1^2 A_{1m} + \beta_1 B_{1m}\right] I_2(m) \exp\left[-\frac{1}{2}S_{m3}(\sigma_1^2 - \sigma_2^2)\right] \sigma_1^{-m} \sigma_2^{-(n-m)} h^{-1}(X) \end{aligned}$$

La distribution a posteriori marginale du point β_2

$$\begin{aligned} g(\beta_1 | X) &= \sum_{m=1}^{n-1} \int_{-\infty}^{+\infty} g(\beta_1, \beta_2, m | X) d\beta_1 \\ &= K_3 \sum_{m=1}^{n-1} \exp\left[-\frac{1}{2}\beta_2^2 A_{2m} + \beta_2 B_{2m}\right] I_1(m) \exp\left[-\frac{1}{2}S_{m3}(\sigma_1^2 - \sigma_2^2)\right] \sigma_1^{-m} \sigma_2^{-(n-m)} h^{-1}(X) \end{aligned}$$

La distribution a posteriori marginale du point m :

$$\begin{aligned}
 g(m | X) &= \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} g(\beta_1, \beta_2, m | X) d\beta_1 d\beta_2 \\
 &= \left[\int_{-\infty}^{+\infty} \exp\left(-\frac{1}{2}\beta_1^2 A_{1m} + \beta_1 B_{1m}\right) d\beta_1 \int_{-\infty}^{+\infty} \exp\left(-\frac{1}{2}\beta_2^2 A_{2m} + \beta_2 B_{2m}\right) d\beta_2 \right. \\
 &\quad \left. \exp\left[-\frac{1}{2}S_{m3}(\sigma_1^{-2} - \sigma_2^{-2})\right] \sigma_1^{-m} \sigma_2^{-(n-m)} \right] \left[\sum_{m=1}^{n-1} T_1(m) \right]^{-1}
 \end{aligned}$$

D'où

$$g(m | X) = \frac{T_1(m)}{\sum_{m=1}^{n-1} T_1(m)}$$

avec :

$$T_1(m) = I_1(m)I_2(m)\exp\left[-\frac{1}{2}S_{m3}(\sigma_1^{-2} - \sigma_2^{-2})\right] \sigma_1^{-m} \sigma_2^{-(n-m)}.$$

3.4 Estimateur Bayésien du point de rupture sous différentes fonctions coûts

3.4.1 La fonction coût quadratique

La fonction coût quadratique est définie par :

$$L_1(\alpha, d) = (\alpha - d)^2$$

L'estimateur de Bayes m^* associé à la distribution marginale $g(m | X)$ sous la fonction coût quadratique est la moyenne a posteriori de m .

L'estimateur de Bayes est :

$$m^* = E(m) = \sum_{m=1}^{n-1} m g(m | X)$$

$$m^* = \frac{\sum_{m=1}^{n-1} m T_1(m)}{\sum_{m=1}^{n-1} T_1(m)}$$

$T_1(m)$ définie précédemment.

3.4.2 La fonction coût absolu

La fonction coût absolu est la fonction définie par :

$$L_2(\alpha, d) = |\alpha - d|$$

L'estimateur de Bayes de m associée à la distribution marginale $g(m | X)$ sous la fonction coût absolu est la médiane de la distribution a posteriori.

3.4.3 La fonction coût 0-1

La fonction coût 0-1 est la fonction définie par :

$$\begin{cases} 0 & \text{si } |\alpha - d| < \epsilon, \quad \epsilon > 0 \\ 1 & ; \text{ sinon.} \end{cases}$$

L'estimateur de Bayes associée à la distribution marginale $g(m | X)$ sous la fonction coût 0-1 est le mode de la distribution a posteriori.

3.5 Analyse Bayésienne en utilisant des lois a priori non informatives

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

Considérons que les lois a priori non informatives de β_1 et β_2 sont des lois normales standardisées et le point de rupture m suit une loi uniforme.

$$\beta_1 \sim N(0, 1)$$

$$\beta_2 \sim N(0, 1)$$

$$m \sim U[1, n - 1]$$

3.5.1 la distribution a priori jointe

de β_1, β_2 et de m sont données par :

$$g_2(\beta_1, \beta_2, m) = [2\pi(n - 1)]^{-1} e^{-\frac{1}{2}\beta_1^2 - \frac{1}{2}\beta_2^2}$$

3.5.2 La distribution a posteriori jointe de β_1, β_2 et de m

$$g_2(\beta_1, \beta_2, m | X) = K_4 \exp\left[-\frac{1}{2}\beta_1^2 A_{3m} + \beta_1 B_{3m}\right] \exp\left[-\frac{1}{2}\beta_2^2 A_{4m} + \beta_2 B_{4m}\right] \\ \exp\left[-\frac{1}{2}S_{m3}(\sigma_1^{-2} - \sigma_2^{-2})\sigma_1^{-m}\sigma_2^{-(n-m)}\right] h_2^{-1}(X)$$

où : $h_2(X)$ est la densité marginale de X

$$h_2(X) = K_4 \sum_{m=1}^{n-1} \int_{-\infty}^{+\infty} \exp\left[-\frac{1}{2}\beta_1^2 A_{3m} + \beta_1 B_{3m}\right] \partial\beta_1 \int_{-\infty}^{+\infty} \exp\left[-\frac{1}{2}\beta_2^2 A_{4m} + \beta_2 B_{4m}\right] \partial\beta_2 \\ \exp\left[-\frac{1}{2}S_{m3}(\sigma_1^{-2} - \sigma_2^{-2})\sigma_1^{-m}\sigma_2^{-(n-m)}\right]$$

La densité a posteriori marginale de β_1

$$g_2(\beta_1 | X) = K_4 \sum_{m=1}^{n-1} \exp[-\frac{1}{2}\beta_1^2 A_{3m} + \beta_1 B_{3m}] I_4(m) \exp[-\frac{1}{2}S_{m3}(\sigma_1^{-2} - \sigma_2^{-2})\sigma_1^{-m}\sigma_2^{-(n-m)}] h_2^{-1}(X).$$

La densité a posteriori marginale de β_2

$$g_2(\beta_2 | X) = K_4 \sum_{m=1}^{n-1} \exp[-\frac{1}{2}\beta_2^2 A_{4m} + \beta_2 B_{4m}] I_3(m) \exp[-\frac{1}{2}S_{m3}(\sigma_1^{-2} - \sigma_2^{-2})\sigma_1^{-m}\sigma_2^{-(n-m)}] h_2^{-1}(X).$$

La densité a posteriori marginale dans le point de rupture m

$$g_2(m | X) = \frac{T_2(m)}{\sum_{m=1}^{n-1} T_2(m)}$$

Où :

$$K_4 = K_2 [2\pi(n-1)] e^{-\frac{S_{n3}}{2}\sigma_2^{-2}}$$

$$A_{3m} = \frac{S_{m1}}{\sigma_1^2} + 1, \quad A_{3m} > 0, \quad B_{3m} = \frac{S_{m2}}{\sigma_1^2},$$

$$A_{4m} = \frac{S_{n1} - S_{m1}}{\sigma_2^2} + 1, \quad A_{4m} > 0, \quad B_{4m} = \frac{S_{n2} - S_{m2}}{\sigma_2^2}.$$

$$I_3(m) = \int_{-\infty}^{+\infty} \exp[-\frac{1}{2}\beta_1^2 A_{3m} + \beta_1 B_{3m}] \partial\beta_1$$

$$I_4(m) = \int_{-\infty}^{+\infty} \exp[-\frac{1}{2}\beta_2^2 A_{4m} + \beta_2 B_{4m}] \partial\beta_2$$

$$T_2(m) = I_3(m) I_4(m) \exp[-\frac{1}{2}S_{m3}(\sigma_1^{-2} - \sigma_2^{-2})\sigma_1^{-m}\sigma_2^{-(n-m)}]$$

3.6 Simulation

Nous allons généré 20 observations issue d'un processus autoregressif d'ordre 1 ($AR(1)$), donnée par le modèle (3.1)

Les valeurs des X_i et des ϵ_i retrouvés pour $i = \overline{1:20}$, sont donnés dans le tableau suivant :

Table-1 : Les observations générés à partir d'un processus autorégressif d'ordre un ($AR(1)$), donné par le modèle (3.1)

I	ϵ_i	X_i	m	ϵ_i	X_i
1	-0.2194347	-0.2194347	11	-3.2479663	-3.4669360
2	0.4048758	0.3829323	12	-3.2668835	-4.3069643
3	1.7709775	1.8092708	13	-7.6887946	-8.9808838
4	-1.8790781	-1.6981511	14	8.5584941	5.8642290
5	1.4848049	1.3149898	15	-0.7197613	1.0395073
6	1.5994697	1.3149898	16	3.8939070	4.2057592
7	-0.6150157	-0.4419189	17	-6.9634980	-5.7017703
8	0.5532121	0.5090202	18	0.1475997	-1.5629314
9	0.1723484	0.2232504	19	-0.3870226	-0.8559020
10	-0.7522240	-0.7298990	20	0.8666289	0.6098583

Les valeurs simulées de la distribution a posteriori marginale du point de rupture m sont données par le tableau suivant :

Table 2 : Les distributions a posteriori marginales de m

m	$g_1(m X)$	$g_2(m X)$	m	$g_1(m X)$	$g_2(m X)$
1	5.076146e-5	9.550951e-05	10	4.639798e-01	4.842691e-01
2	9.941465e-5	1.840302e-04	11	3.467914e-01	3.158194e-01
3	1.569535e-04	3.400712e-04	12	1.830399e-12	4.128126e-11
4	5.656141e-04	8.463237e-04	13	2.319560e-17	2.631310e-17
5	2.016433e-03	2.675474e-03	14	3.551508e-17	3.985351e-17
6	8.038483e-03	1.054646e-02	15	2.225317e-21	2.476917e-21
7	1.133415e-02	1.516055e-02	16	7.060802e-23	9.487329e-23
8	3.851864e-02	3.899246e-02	17	6.263011e-28	1.239690e-27
9	1.284484e-01	1.310707e-01	18	2.372646e-27	5.865434e-27
			19	1.010043e-26	3.181201e-26

$g_1(m | X)$ correspond au cas où les loi a priori sont informatives, (c'est à dire : Une loi normale de moyenne μ_i et de variance a_i , pour les β_i avec $i = 1, 2$. (les valeurs des μ_i et a_i sont données précédement .

Et $g_2(m | X)$ correspond au cas où elles sont non informatives, (des lois normales standares $N(0, 1)$ pour β_1 et β_2).

Remarque : Nous remarquons que le cas non informative est un cas particulier du cas informative avec : $\mu_1 = \mu_2 = 0$ et $a_1 = a_2 = 1$. Donc pour la simulation, nous avons gardé le même programme, (donné en annexe) avec : $\mu_1 = \mu_2 = 0$ et $a_1 = a_2 = 1$.

Calcul des estimateurs Bayésien du point de rupture m :

Nous allons utiliser les différentes fonctions coût usuelles : la quadratique, la coût 0 – 1 et l'absolue. Pour cela, nous allons calculer la moyenne, la médiane et le mode a posteriori.

La moyenne a posteriori

Dans le cas informatif, la moyenne a posteriori est donnée par :

$$m^* = \sum_{m=1}^{n-1} m g_1(m | X) = 9.694367$$

Dans le cas non informatif, la moyenne a posteriori est donnée par :

$$m^{**} = \sum_{m=1}^{n-1} m g_2(m | X) = 9.630973$$

Le mode a posteriori

Nous allons calculer la plus grande valeur de $g_1(x)$, d'après les résultats le $\max(g_1(x)) = 0.4639798$. La valeur correspondante à ce max est égale à 10.

Dans le cas non informatif, le $\max(g_2(x)) = 0.4842691$. La valeur correspondante à ce max est égale à 10.

Donc dans les deux cas, le mode a posteriori est égale à 10.

Pour mieux illustrer ce résultats, nous avons tracé les graphes des distributions a posteriori marginales de m , dans les deux cas : informatives et non informatives.

Les graphes sont donnés dans les figures suivantes :

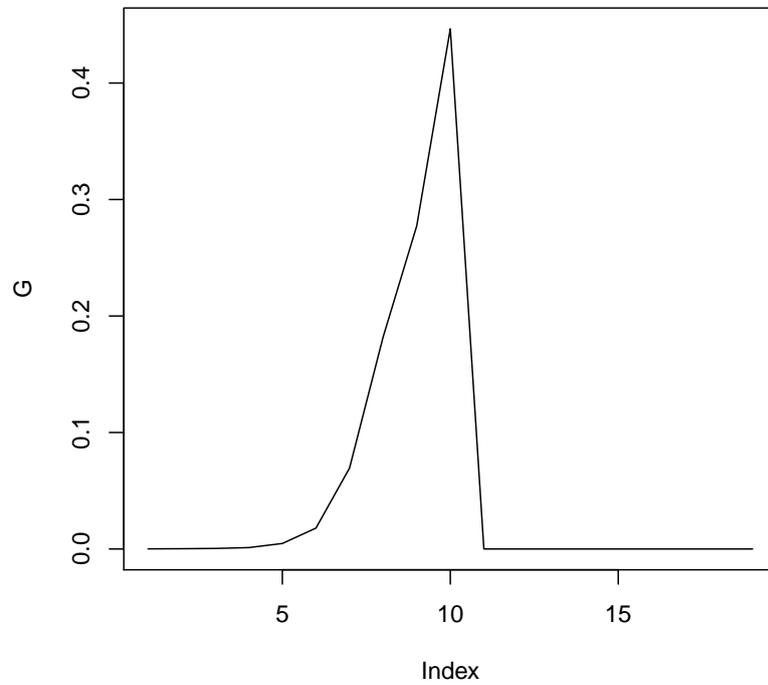


Figure3-1 : graphe de la distribution a posteriori de m pour $n = 20$, $m = 10$ dans le cas informative.

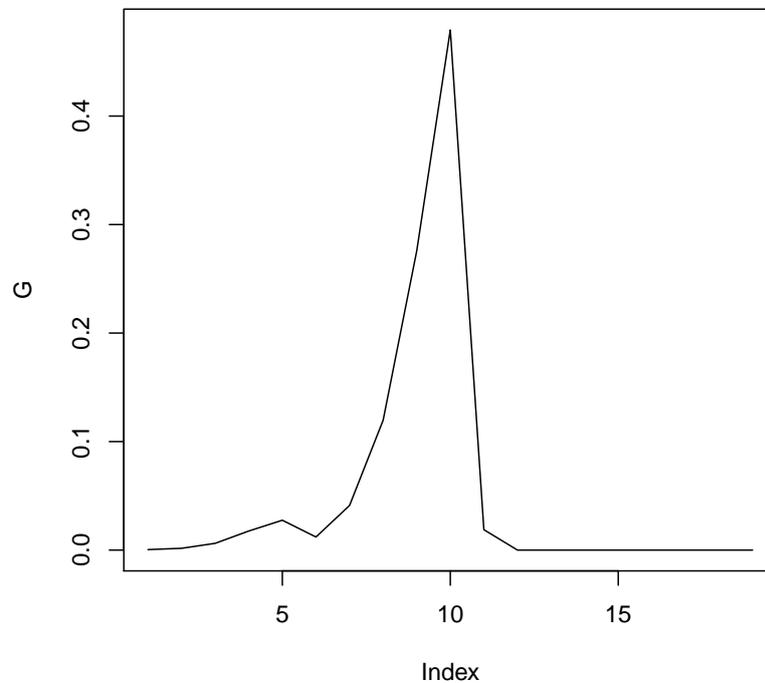


Figure3-2 : graphe de la distribution a posteriori de m pour $n = 20$, $m = 10$ dans le cas non informative.

La mediane a posteriori

Afin de calculer la mediane a posteriori, nous allons calculé la distribution a posteriori cumulée de m . Elles sont donnée dans le tableau suivant :

Table-3 : La distribution a posteriori cumulée de m

m	$F_1(m X)$	$F_2(m X)$	m	$F_1(m X)$	$F_2(m X)$
1	7.305731e-05	0.0001099834	10	5.787293e-01	0.9999770846
2	2.312670e-04	0.0003309571	11	7.648541e-01	0.9999971564
3	7.029119e-04	0.0007972239	12	8.110281e-01	0.9999999983
4	2.493687e-03	0.0019893119	13	9.754911e-01	0.9999999987
5	5.119849e-03	0.0066991137	14	9.845896e-01	1.0000000000
6	1.408152e-02	0.0246346355	15	9.998782e-01	1.0000000000
7	3.609041e-02	0.0941108903	16	9.999573e-01	1.0000000000
8	1.134532e-01	0.2761245674	17	9.999729e-01	1.0000000000
9	4.096029e-01	0.5534437323	18	1.000000e+00	1.0000000000
			19	1.000000e+00	1.0000000000

D'après le tableau des distribution a posteriori cumulée de m :

La médiane a posteriori est égale à 10 dans le cas **informatif**.

La médiane a posteriori est égale à 9 dans le cas **non informatif**.

nous remarquons que les trois estimateurs (sous les fonction couts quadratique 0-1 absolue) sont très bien représentés.

Position du point de rupture

pour mieux illustrer cette méthode d'estimation du point de rupture, nous avons fait varier sa position. Nous considérons deux cas : (proche de 1 et proche de n). Pour cela, nous avons pris dans un premier temps $m = 5$, puis $m = 15$.

Nous avons gardé le même programme en faisant varier la valeur de m .

Les résultats retrouvés sont :

Pour $m = 5$

Les distributions a posteriori du point de rupture dans ce cas où $m = 5$, sont données respectivement, dans les deux cas (lois informatives par $g_1(x)$, et non informatives par

$g_2(x)$, dans le tableau suivant :

Table-3 : Les distributions a posteriori marginales pour $n = 20$ et $m = 5$

m	$g_1(m X)$	$g_2(m X)$	m	$g_1(m X)$	$g_2(m X)$
1	4.654098e-03	1.334467e-02	10	2.140326e-08	1.154512e-09
2	9.117661e-03	2.641779e-02	11	2.326823e-08	3.783143e-09
3	1.441604e-02	4.099613e-02	12	7.247897e-22	1.678843e-09
4	5.191230e-02	1.358284e-01	13	1.228803e-24	6.059539e-10
5	1.849654e-01	4.998889e-01	14	2.362906e-24	1.681778e-09
6	7.349340e-01	1.768950e-01	15	1.505736e-28	5.346877e-09
7	1.394277e-07	1.012465e-01	16	3.825638e-30	1.669653e-08
8	2.861420e-07	5.382570e-03	17	9.637890e-36	4.484241e-10
9	7.386635e-08	2.923750e-10	18	2.636415e-35	2.462768e-10
			19	1.034717e-34	1.698310e-10

Dans le cas des loi a priori **informatives**, la **moyenne a posteriori** est égale à :
5.243266

Dans le cas des loi a priori **non informatives**, la **moyenne a posteriori** est égale à :
4.680127

Le mode a posteriori

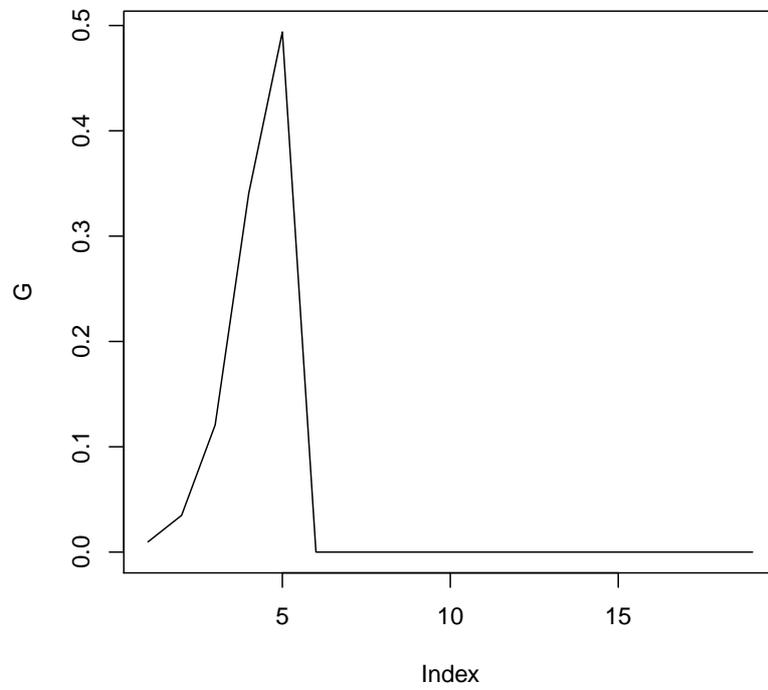


Figure3-3 : graphe de la distribution a posteriori de m pour $n = 20$, $m = 5$ dans le cas non informative.

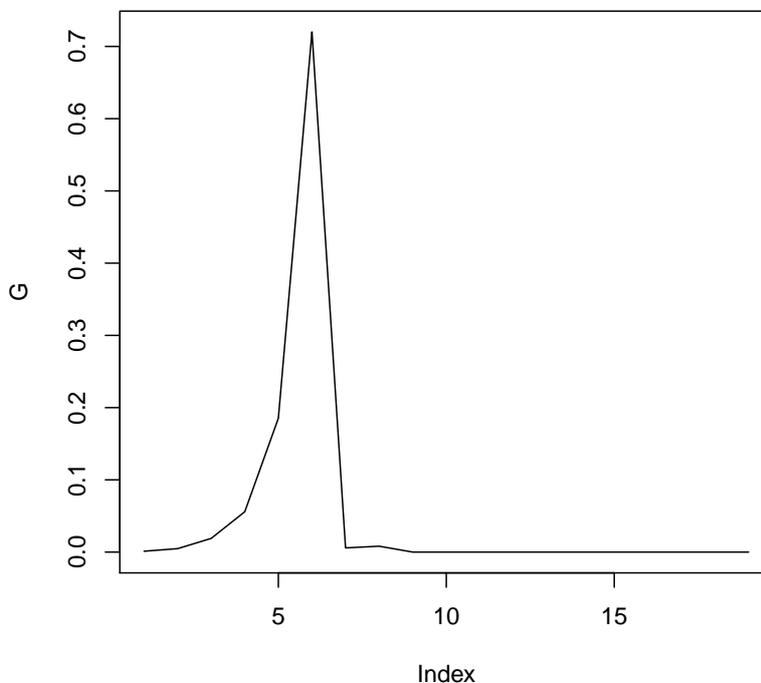


Figure3-4 : graphe de la distribution a posteriori de m pour $n = 20$, $m = 5$ dans le cas non informative.

Dans le cas des loi a priori **informatives**, la **mediane** est égale à : 5.5

Dans le cas des loi a priori **non informatives**, la **mediane** est égale à : 4.5

Pour $m = 15$

Les distributions a posteriori du point de rupture dans ce cas où $m = 15$, sont données respectivement dans le tableau suivant, dans les deux cas (lois informatives par $g_1(x)$, et non informatives par $g_2(x)$).

Table 2 : La distribution a posteriori pour $n = 20$ et $m = 15$

m	$g_1(m X)$	$g_2(m X)$	m	$g_1(m X)$	$g_2(m X)$
1	1.787395e-06	5.788006e-08	10	1.641647e-02	3.389874e-04
2	3.499559e-06	1.173646e-07	11	5.624842e-02	1.343150e-03
3	5.521208e-06	1.822275e-07	12	2.414952e-02	4.651781e-03
4	1.993506e-05	6.100418e-07	13	7.296956e-02	1.533214e-02
5	7.112011e-05	2.258682e-06	14	2.679372e-01	5.898238e-02
6	2.835483e-04	7.196126e-06	15	5.469261e-01	2.344811e-01
7	3.997590e-04	2.284977e-05	16	8.650186e-03	5.880955e-01
8	1.362960e-03	3.818931e-05	17	1.544803e-06	1.451246e-02
9	4.544704e-03	8.689823e-05	18	1.923724e-06	2.778053e-02
			19	6.244251e-06	5.432360e-02

Dans le cas des loi a priori **informatives**, la **moyenne a posteriori** est égale à :
13.8067

Dans le cas des loi a priori **non informatives**, la **moyenne a posteriori** est égale à :
15.44106

Le mode a posteriori est : 15

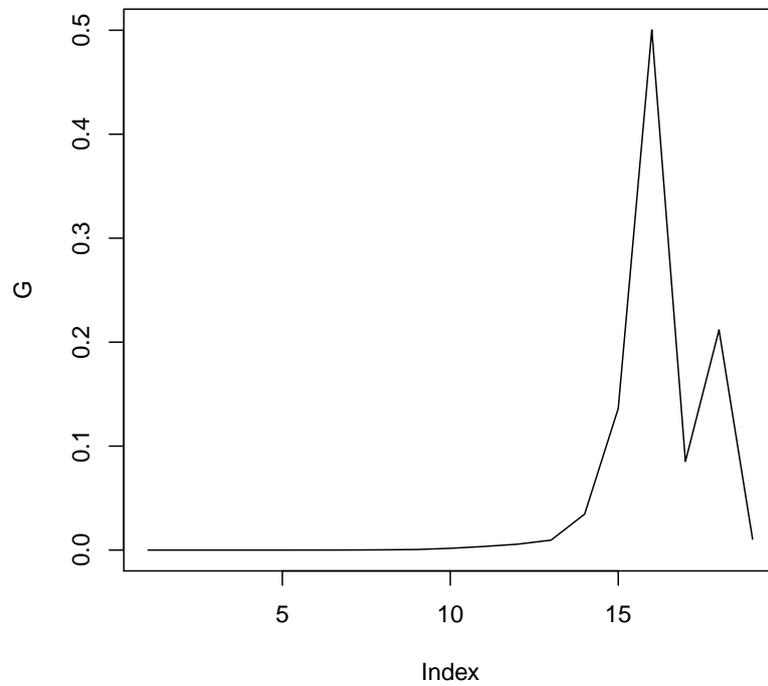


Figure3-5 : graphe de la distribution a posteriori de m pour $n = 20$, $m = 15$ dans le cas non informative.

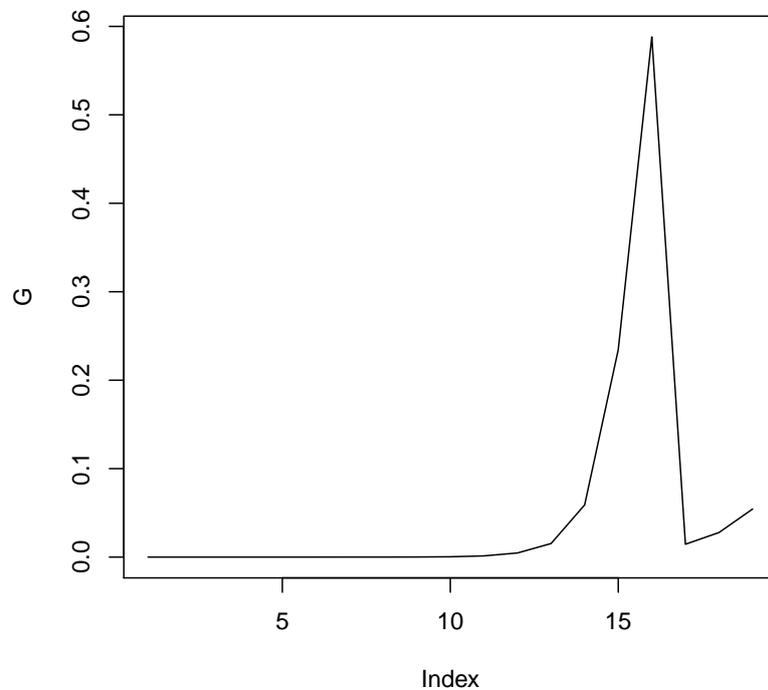


Figure3-6 : graphe de la distribution a posteriori de m pour $n = 20$, $m = 15$ dans le cas non informative.

Dans le cas des loi a priori **informatives**, la **mediane** est égale à : 14.5

Dans le cas des loi a priori **non informatives**, la **mediane** est égale à : 15.5

Conclusion :

D'après les résultats de simulation retrouvés, nous remarquons que la méthode Bayésienne estime bien la valeur du point de rupture, dans les deux cas où les lois a priori sont informatives ou pas.

Nous remarquons également que peut importe le choix de la fonction coût (quadratique,

coût 0-1 ou absolue), nous retrouvons de très bons estimateurs Bayésien.

Et enfin, la dernière simulation nous montre que la position du point de rupture n'a aucune influence sur cette méthode d'estimation. Les résultats sont toujours satisfaisants.

Conclusion et perspectives

Dans ce travail nous avons élaboré les séries chronologiques où nous avons vu les différents modèles et quelques notions de base dans le cas où les erreurs sont stationnaires et non stationnaires.

En suites nous avons étudié les problèmes de rupture et quelques méthodes de détection de rupture parmi ces méthodes, nous retrouvons : les tests statistiques qui détectent la présence d'une rupture, et la méthode Bayésienne pour l'estimation de ce point.

La méthodologie Bayésienne apporte une grande souplesse dans les méthodologies statistiques, elle a été largement sollicitée dans les problèmes de détection de rupture.

Enfin, nous avons traitées dans la partie simulation l'estimation Bayésienne d'un point de rupture dans un processus autorégressif d'ordre un ($AR(1)$), et nous avons conclu que les trois estimateurs (moyenne, mode et médiane) sont des bons estimateurs.

Il serait intéressant d'étendre ce travail sous les différents types de dépendances. Aussi peut être est il possible d'améliorer l'estimation en choisissant des lois a priori plus générales.