

Université de Montréal

Impact de la taille de la partition de
l'espace-paramètre sur les résultats des tests
d'hypothèses multiples sous différentes fonctions
de perte

par

11592956

Étienne Chassé St-Laurent

Département de mathématiques et de statistique
Faculté des arts et des sciences

Mémoire présenté à la Faculté des études supérieures
en vue de l'obtention du grade de
Maître ès sciences (M.Sc.)
en Statistique

octobre 2004



QA

3

U54

2004

V. 015

Direction des bibliothèques

AVIS

L'auteur a autorisé l'Université de Montréal à reproduire et diffuser, en totalité ou en partie, par quelque moyen que ce soit et sur quelque support que ce soit, et exclusivement à des fins non lucratives d'enseignement et de recherche, des copies de ce mémoire ou de cette thèse.

L'auteur et les coauteurs le cas échéant conservent la propriété du droit d'auteur et des droits moraux qui protègent ce document. Ni la thèse ou le mémoire, ni des extraits substantiels de ce document, ne doivent être imprimés ou autrement reproduits sans l'autorisation de l'auteur.

Afin de se conformer à la Loi canadienne sur la protection des renseignements personnels, quelques formulaires secondaires, coordonnées ou signatures intégrées au texte ont pu être enlevés de ce document. Bien que cela ait pu affecter la pagination, il n'y a aucun contenu manquant.

NOTICE

The author of this thesis or dissertation has granted a nonexclusive license allowing Université de Montréal to reproduce and publish the document, in part or in whole, and in any format, solely for noncommercial educational and research purposes.

The author and co-authors if applicable retain copyright ownership and moral rights in this document. Neither the whole thesis or dissertation, nor substantial extracts from it, may be printed or otherwise reproduced without the author's permission.

In compliance with the Canadian Privacy Act some supporting forms, contact information or signatures may have been removed from the document. While this may affect the document page count, it does not represent any loss of content from the document.

Université de Montréal

Faculté des études supérieures

Ce mémoire intitulé

Impact de la taille de la partition de
l'espace-paramètre sur les résultats des tests
d'hypothèses multiples sous différentes fonctions
de perte

présenté par

Étienne Chassé St-Laurent

a été évalué par un jury composé des personnes suivantes :

Charles Dugas

(président-rapporteur)

Jean-François Angers

(directeur de recherche)

Louis Doray

(co-directeur)

Martin Bilodeau

(membre du jury)

Mémoire accepté le:

Le 6 octobre 2004

SOMMAIRE

Dans ce mémoire, nous nous intéressons à des tests d'hypothèses multiples basées sur une partition de l'espace-paramètre ainsi qu'aux relations entre ces tests et les estimateurs ponctuels utilisant les fonctions de perte ponctuelles correspondantes, dans un contexte bayésien. Les partitions sont caractérisées par le fait que nous divisons l'espace-paramètre en intervalles de largeur égale et que chaque intervalle est assigné à une hypothèse spécifique.

Nous déterminons la décision de Bayes pour six fonctions de perte différentes utilisées fréquemment, puis nous appliquons ces résultats à trois types de distributions caractérisées par la symétrie et le nombre de modes de leur densité. Finalement, nous trouvons la décision de Bayes lorsque la largeur des intervalles tend vers 0, et nous comparons avec les estimateurs ponctuels des fonctions de perte correspondantes. Nous finissons avec des simulations et des illustrations ainsi qu'un exemple provenant de données réelles.

SUMMARY

We present our results on multiple hypothesis tests based on a partition of the parameter-space and the comparison of these results with point estimation of the parameter using their corresponding loss functions, in a Bayesian context. Our partition is such that the parameter-space is divided into equal-length intervals which are assigned in a one-to-one relation to a specific hypothesis.

We find the Bayes' decision for six different usual loss functions, then we apply them to three types of distributions characterized by their symmetry and their number of modes. Finally, we find the limiting behavior of these results when the length of the intervals tends to 0, and we compare them with the point estimates found with the corresponding loss functions. We conclude with simulations and illustrations of the results and an example based on real life data.

MOTS-CLÉS

Statistique bayésienne ; hypothèse précise ; hypothèse ponctuelle ; estimation ponctuelle ; fonction de perte ; LINEX ; tests d'hypothèses multiples ; partition- ϵ ; simulation ; données d'assurance-incendie.

KEYWORDS

Bayesian statistics; precise hypothesis; point hypothesis; point estimation; loss function; LINEX; multiple hypothesis test; ϵ -partition; simulation; fire insurance data.

TABLE DES MATIÈRES

Sommaire	iii
Summary	iv
Mots-clés	v
Keywords	vi
Liste des figures	xi
Liste des tableaux	xii
Remerciements	xiii
Introduction	3
Introduction	3
Chapitre 1. Tests bayésiens d’hypothèses multiples	5
1.1. Statistique bayésienne, tests d’hypothèses multiples et estimation ponctuelle.....	5
1.2. Fonctions de perte à valeurs constantes.....	7
1.2.1. Fonctions de perte 0-1 et 0-k (L_1 et L_2).....	7
1.2.2. Fonction de perte L_3	8
1.2.3. Fonction de perte L_4	9
1.3. Fonctions de perte à valeurs continues.....	11
1.3.1. Fonction de perte quadratique (L_5).....	11
1.3.2. Fonction de perte LINEX (intégrée) (L_6).....	12

1.4. Conclusion.....	19
Chapitre 2. Distributions et résultats spécifiques.....	20
2.1. Distribution unimodale symétrique.....	21
2.1.1. Caractéristiques.....	21
2.1.2. Résultats.....	21
2.1.2.1. Fonctions de perte 0-1 et 0-k (L_1 et L_2).....	21
2.1.2.2. Fonction de perte L_3	23
2.1.2.3. Fonction de perte L_4	24
2.1.2.4. Fonction de perte quadratique (L_5).....	25
2.1.2.5. Fonction de perte LINEX (L_6).....	25
2.2. Distribution unimodale asymétrique.....	26
2.2.1. Caractéristiques.....	26
2.2.2. Résultats.....	26
2.2.2.1. Fonctions de perte 0 – 1 et 0 – k (L_1 et L_2).....	26
2.2.2.2. Fonction de perte L_3	28
2.2.2.3. Fonction de perte L_4	28
2.2.2.4. Fonction de perte quadratique (L_5).....	29
2.2.2.5. Fonction de perte LINEX (L_6).....	29
2.3. Distribution bimodale.....	29
2.3.1. Caractéristiques.....	29
2.3.2. Résultats.....	30
2.3.2.1. Fonctions de perte 0-1 et 0-k (L_1 et L_2).....	30
2.3.2.2. Fonction de perte L_3	31
2.3.2.3. Fonction de perte L_4	31
2.3.2.4. Fonction de perte quadratique (L_5).....	31
2.3.2.5. Fonction de perte LINEX (L_6).....	32
2.4. Distribution multimodale.....	32

2.4.1. Caractéristiques.....	32
2.4.2. Résultat pour les fonctions de perte 0-1 et 0-k (L_1 et L_2).....	32
2.5. Conclusion.....	36
Chapitre 3. Comparaison avec l'estimation ponctuelle, simulations et exemple	38
3.1. Comparaison avec l'estimation ponctuelle	38
3.1.1. Fonctions de perte à valeurs constantes.....	38
3.1.1.1. Fonctions de perte 0-1 et 0-k (L_1 et L_2).....	38
3.1.1.2. Fonction de perte L_3	42
3.1.1.3. Fonction de perte L_4	43
3.1.2. Fonctions de perte à valeurs continues.....	44
3.1.2.1. Fonction de perte quadratique (L_5).....	44
3.1.2.2. Fonction de perte LINEX (L_6).....	44
3.2. Simulations.....	48
3.2.1. Accord entre estimation ponctuelle et test d'hypothèses multiples	48
3.2.1.1. Fonction de perte 0-1 (L_1).....	49
3.2.1.2. Fonctions de perte L_3 et L_4	49
3.2.1.3. Fonction de perte quadratique (L_5).....	49
3.2.1.4. Fonction de perte LINEX (L_6).....	49
3.2.2. Simulation pour la fonction de perte LINEX (L_6).....	51
3.2.3. Illustration des résultats pour les trois types de distribution....	53
3.3. Exemple.....	67
3.4. Conclusion.....	70
Conclusion.....	71
Annexe A. Théorème sur les distributions multimodales	A-i

Bibliographie A-i

LISTE DES FIGURES

1.3.1	Illustration de la perte pour la fonction de perte LINEX ponctuelle ..	13
3.2.1	Probabilité <i>a posteriori</i> , constantes pour L_3 et L_4 , pertes espérées <i>a posteriori</i> pour L_5 et L_6 pour chacun des 50 intervalles, pour la distribution unimodale symétrique	61
3.2.2	Probabilité <i>a posteriori</i> , constantes pour L_3 et L_4 , pertes espérées <i>a posteriori</i> pour L_5 et L_6 pour chacun des 100 intervalles, pour la distribution unimodale symétrique	62
3.2.3	Probabilité <i>a posteriori</i> , constantes pour L_3 et L_4 , pertes espérées <i>a posteriori</i> pour L_5 et L_6 pour chacun des 50 intervalles, pour la distribution unimodale asymétrique	63
3.2.4	Probabilité <i>a posteriori</i> , constantes pour L_3 et L_4 , pertes espérées <i>a posteriori</i> pour L_5 et L_6 pour chacun des 100 intervalles, pour la distribution unimodale asymétrique	64
3.2.5	Probabilité <i>a posteriori</i> , constantes pour L_3 et L_4 , pertes espérées <i>a posteriori</i> pour L_5 et L_6 pour chacun des 50 intervalles, pour la distribution bimodale	65
3.2.6	Probabilité <i>a posteriori</i> , constantes pour L_3 et L_4 , pertes espérées <i>a posteriori</i> pour L_5 et L_6 pour chacun des 100 intervalles, pour la distribution bimodale	66
3.3.1	Comparaison entre la distribution du logarithme des données réelles et le modèle	67

LISTE DES TABLEAUX

3.2.1	Nombre de fois où l'estimateur ponctuel est dans l'intervalle retenu par le test à hypothèses multiples, pour les trois types de distributions, sur 10 000 simulations.	50
3.2.2	Nombre de fois où l'estimateur ponctuel de la fonction de perte LINEX est dans l'intervalle retenu par le test à hypothèses multiples, pour les trois types de distribution, en utilisant $\alpha = 2$	52
3.2.3	Illustration des pertes espérées pour les 6 fonctions de perte, avec 500 données simulées provenant d'une normale(0, 500) et 20 intervalles. ...	55
3.2.4	Illustration des pertes espérées pour les 6 fonctions de perte, avec 500 données simulées provenant d'une gamma(20, θ) et 20 intervalles.	57
3.2.5	Illustration des pertes espérées pour les 6 fonctions de perte, avec 500 données simulées provenant d'un mélange de deux normales et 20 intervalles.	59
3.3.1	Illustration des pertes espérées pour les 6 fonctions de perte, avec le jeu de données réelles et 20 intervalles.	68

REMERCIEMENTS

Je commence par remercier mon directeur de recherche Jean-François, pour m'avoir introduit aux statistiques, aux statistiques bayésiennes et m'avoir suggéré mon projet de recherche. J'aimerais aussi le remercier pour son support et sa disponibilité malgré son horaire surchargé. Ses suggestions m'ont été très utiles tout au long de ma maîtrise. Je voudrais aussi remercier mon codirecteur de recherche Louis Doray pour les données utilisées dans mon exemple ainsi que pour m'avoir introduit à la fonction de perte LINEX, qui est devenue une des parties les plus intéressantes de mon mémoire.

Je veux aussi remercier toute ma famille pour leur support tout au long de mes études, mon père, ma mère ainsi que mes deux frères. Je voudrais remercier aussi tous mes amis de maîtrise pour les bons moments que nous avons passés ensemble : Élyse, Nadine, Anne, Géraldine, Alain, René, Martine, Geneviève, Javier, Louis-François, Alice et James. Un merci spécial à Alice pour m'avoir laissé m'inspirer de son mémoire pour le mien ainsi qu'à Nadine pour ses conseils esthétiques.

J'aimerais finalement remercier tous mes professeurs au département ainsi que tout le personnel de soutien. Je veux aussi remercier le CRSNG pour la bourse de maîtrise qui m'a permis de faire mes études.

INTRODUCTION

Dans ce mémoire nous nous sommes penchés sur les tests d'hypothèses multiples où chaque hypothèse correspond à un des intervalles d'une partition de l'espace-paramètre en intervalles de largeurs égales. Notre motivation vient de la discussion après l'exemple 5 de l'article de Evans (1997) où il affirme que pour l'approche de surprise relative, il n'y a aucune différence entre les tests d'hypothèses ponctuelles ou d'hypothèses sur des intervalles disjoints. Nous avons décidé d'explorer ce sujet dans le cadre de l'utilisation de fonctions de perte.

Dans le premier chapitre, nous présentons un bref rappel (ou une brève introduction, selon le cas) à la statistique bayésienne. Ensuite, nous présentons les six fonctions de perte que nous avons sélectionnées ainsi que quelques résultats préliminaires pour chacune de ces fonctions de perte.

Au deuxième chapitre, nous nous sommes intéressés aux résultats des tests d'hypothèses multiples avec ces fonctions de perte dans le cas de trois types de distributions différents, caractérisés par la symétrie de la densité et le nombre de modes. En plus, pour la fonction de perte 0-1 (la plus simple), nous avons exploré un type de distribution plus général où le nombre de modes est fini. Il est à noter que nos résultats pour la fonction de perte 0-1 sont valides seulement si la densité tend vers 0 aux deux extrémités du support et qu'elle est continue.

Au troisième chapitre, nous présentons une comparaison des résultats des tests d'hypothèses multiples lorsque nous faisons converger la largeur des intervalles vers 0 avec les résultats de l'estimation ponctuelle. Ensuite, nous présentons des

simulations pour illustrer nos résultats précédents ainsi qu'un exemple utilisant des données réelles, provenant d'un article de Ramlau-Hansen (1988).

À l'annexe, nous présentons un théorème très général pour le nombre d'intervalles qui peuvent avoir la probabilité *a posteriori* maximale pour une distribution multimodale dont la densité est continue et tend vers 0 aux deux extrémités de son support.

INTRODUCTION

Dans ce mémoire nous nous sommes penchés sur les tests d'hypothèses multiples où chaque hypothèse correspond à un des intervalles d'une partition de l'espace-paramètre en intervalles de largeurs égales. Notre motivation vient de la discussion après l'exemple 5 de l'article de Evans (1997) où il affirme que pour l'approche de surprise relative, il n'y a aucune différence entre les tests d'hypothèses ponctuelles ou d'hypothèses sur des intervalles disjoints. Nous avons décidé d'explorer ce sujet dans le cadre de l'utilisation de fonctions de perte.

Dans le premier chapitre, nous présentons un bref rappel (ou une brève introduction, selon le cas) à la statistique bayésienne. Ensuite, nous présentons les six fonctions de perte que nous avons sélectionnées ainsi que quelques résultats préliminaires pour chacune de ces fonctions de perte.

Au deuxième chapitre, nous nous sommes intéressés aux résultats des tests d'hypothèses multiples avec ces fonctions de perte dans le cas de trois types de distributions différents, caractérisés par la symétrie de la densité et le nombre de modes. En plus, pour la fonction de perte 0-1 (la plus simple), nous avons exploré un type de distribution plus général où le nombre de modes est fini. Il est à noter que nos résultats pour la fonction de perte 0-1 sont valides seulement si la densité tend vers 0 aux deux extrémités du support et qu'elle est continue.

Au troisième chapitre, nous présentons une comparaison des résultats des tests d'hypothèses multiples lorsque nous faisons converger la largeur des intervalles vers 0 avec les résultats de l'estimation ponctuelle. Ensuite, nous présentons des

simulations pour illustrer nos résultats précédents ainsi qu'un exemple utilisant des données réelles, provenant d'un article de Ramlau-Hansen (1988).

À l'annexe, nous présentons un théorème très général pour le nombre d'intervalles qui peuvent avoir la probabilité *a posteriori* maximale pour une distribution multimodale dont la densité est continue et tend vers 0 aux deux extrémités de son support.

Chapitre 1

TESTS BAYÉSIENS D'HYPOTHÈSES MULTIPLES

Dans ce premier chapitre, nous exposons la théorie derrière les résultats présentés dans ce mémoire ainsi que quelques résultats préliminaires.

Nous commençons par une présentation rapide de la notation et des concepts pertinents au sujet, principalement les tests bayésiens d'hypothèses multiples. Nous ajoutons une courte explication sur l'estimation ponctuelle.

Ensuite, nous présentons les six fonctions de perte qui sont étudiées dans ce mémoire. Pour chacune (sauf les deux premières qui sont traitées ensemble), la section contient l'équation de la fonction de perte suivie de résultats la concernant, incluant l'intervalle qui sera choisi lorsqu'il est possible de le préciser. Il est à noter que les résultats obtenus sont très généraux, étant donné que la distribution *a posteriori* n'est pas spécifiée.

1.1. STATISTIQUE BAYÉSIENNE, TESTS D'HYPOTHÈSES MULTIPLES ET ESTIMATION PONCTUELLE

Étant donné que ce mémoire se situe exclusivement dans un contexte bayésien, il nous a semblé utile de présenter un bref rappel sur la statistique bayésienne en

mettant particulièrement l'accent sur les tests d'hypothèses multiples et l'estimation ponctuelle.

En statistique bayésienne, notre paramètre d'intérêt possède une densité initiale (ou *a priori*) que nous notons $\pi(\theta)$. Après avoir observé un échantillon, noté (X_1, X_2, \dots, X_n) , dont nous connaissons la densité $f(x|\theta)$, $i = 1, 2, \dots, n$, nous pouvons mettre à jour la densité du paramètre à l'aide du théorème de Bayes. Nous obtenons alors comme densité *a posteriori*

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

où Θ est l'espace-paramètre de θ , soit l'ensemble des valeurs possibles pour θ .

Pour tout ce mémoire, nous ne nous attardons pas à la distribution *a priori* du paramètre ou celle des observations. Nous assumons que nous connaissons déjà la distribution *a posteriori*.

Nous devons ensuite définir une fonction de perte. Étant donné que nous ne pouvons pas obtenir des résultats pour toutes les fonctions de perte possibles, nous avons choisi quelques-unes des plus courantes.

Définition 1.1.1. Une fonction de perte, notée $L(\theta, a)$, associe une valeur numérique (la perte comme telle) à une décision, notée a , pour une valeur donnée, notée θ , du paramètre.

Définition 1.1.2. Une fois la fonction de perte définie, nous prenons la décision qui minimise la perte espérée *a posteriori* (par rapport à la distribution *a posteriori* du paramètre), c'est-à-dire $\rho(\theta, a|x) = \mathbb{E}^{\pi(\theta|x)}[L(\theta, a)]$. Cette décision est appelée la décision de Bayes, et elle est notée $\delta^\pi(x)$, avec $\delta^\pi(x) = \arg \min_a \mathbb{E}^{\pi(\theta|x)}[L(\theta, a)]$.

Si nous traitons d'estimation ponctuelle, la décision correspondra à une valeur précise qui minimise la perte espérée *a posteriori*. Cette valeur nous servira d'estimateur ponctuel.

Pour ce qui est d'un test d'hypothèses multiples, nous n'avons qu'à définir

$$H_1 : \theta \in \Theta_1, H_2 : \theta \in \Theta_2, \dots, \bigcup_{i \geq 1} \Theta_i = \Theta, \Theta_i \cap \Theta_j = \emptyset, i \neq j$$

et faire correspondre à chaque hypothèse une décision. Ainsi, nous choisissons l'hypothèse qui correspond à la décision minimisant la perte espérée *a posteriori*. Dans notre cas, nous nous intéressons à des hypothèses de la forme (tous les intervalles ont la même largeur)

$$H_i : \theta \in \Theta_i, \text{ où } \Theta_i = (\theta_0 + (i-1)\varepsilon; \theta_0 + i\varepsilon], i \in \mathbb{Z}, \theta_0 \in \mathbb{R}.$$

1.2. FONCTIONS DE PERTE À VALEURS CONSTANTES

1.2.1. Fonctions de perte 0-1 et 0-k (L_1 et L_2)

Cette première fonction de perte ne nécessitera qu'une brève introduction, étant la plus simple de toutes. Elle existe sous deux formes qui reviennent au même dans notre cas.

La première forme donne une perte de 1 pour une décision erronée et 0 pour la bonne décision. La seconde, légèrement plus générale, remplace le 1 par une constante k (nous assumons que $k > 0$). Ceci ne fait que modifier toutes les pertes par une constante multiplicative et ne change pas nos résultats, qui dépendent de comparaisons. Donc, nos première et deuxième fonctions de perte sont :

$$L_1(\theta, d) = \begin{cases} 0 & \text{si } \theta \in \Theta_d; \\ 1 & \text{sinon;} \end{cases}$$

$$L_2(\theta, d) = \begin{cases} 0 & \text{si } \theta \in \Theta_d; \\ k & \text{sinon. } k > 0. \end{cases}$$

Les résultats étant les mêmes pour ces deux premières fonctions de perte, nous étudierons seulement la première. Trouvons la décision de Bayes qui minimise la perte espérée *a posteriori*, qui est donnée par

$$\rho_1(\theta, a|x) = \mathbb{E}^{\pi(\theta|x)}[L_1(\theta, a)] = 1 - \Pr\{\theta \in \Theta_a|x\}.$$

Donc $\rho_1(\theta, a|x)$ est minimal lorsque $\Pr\{\theta \in \Theta_a|x\}$ est maximal, c'est-à-dire

$$\delta^\pi(x) = \arg \max_{a \in \mathbb{Z}} \Pr\{\theta \in \Theta_a|x\}.$$

1.2.2. Fonction de perte L_3

Une première façon de généraliser L_2 est d'assigner une perte constante spécifique à chaque décision si elle est erronée et une perte nulle si cette décision est la bonne. Nous obtenons alors une troisième fonction de perte,

$$L_3(\theta, d) = \begin{cases} 0 & \text{si } \theta \in \Theta_d; \\ k_d & \text{sinon, } k_d > 0, \forall d \in \mathbb{Z}. \end{cases}$$

Cette fonction de perte est plus flexible que les deux précédentes. En effet, nous pouvons maintenant pénaliser différemment les décisions, selon leur niveau de « risque ». Ceci s'avère préférable dans beaucoup de situations. Il est important de souligner que la perte (s'il y a perte) dépend seulement de la décision qui est prise et non de la vraie valeur de θ . Ainsi, la bonne décision n'influence la perte que si nous la choisissons et dans ce cas, la perte est 0. Sinon, la perte dépend uniquement de la décision d , avec $\theta \notin \Theta_d$ puisque d n'est pas la bonne décision.

Nous cherchons maintenant les caractéristiques de la décision de Bayes pour cette fonction de perte. Nous avons donc

$$\begin{aligned} \rho_3(\theta, a|x) &= \mathbb{E}^{\pi(\theta|x)}[L_3(\theta, a)] \\ &= k_a(1 - \Pr\{\theta \in \Theta_a|x\}) \\ &= k_a(1 - \Pi_a), \end{aligned}$$

où $\Pi_a = \Pr\{\theta \in \Theta_a|x\}$.

Nous préférons une décision a_1 à une autre décision a_2 si et seulement si la perte espérée *a posteriori* de a_1 est inférieure à celle de a_2 . Donc, si et seulement si

$$\rho_3(\theta, a_1|x) < \rho_3(\theta, a_2|x) \Leftrightarrow k_{a_1}(1 - \Pi_{a_1}) < k_{a_2}(1 - \Pi_{a_2}).$$

De ceci nous obtenons trois conditions équivalentes,

- (1) $k_{a_1} < k_{a_2} \left(\frac{1 - \Pi_{a_2}}{1 - \Pi_{a_1}} \right)$;
- (2) $\Pi_{a_1} > 1 - \frac{k_{a_2}}{k_{a_1}}(1 - \Pi_{a_2})$;
- (3) $k_{a_1}\Pi_{a_1} > k_{a_2}\Pi_{a_2} + (k_{a_1} - k_{a_2})$.

D'après la deuxième condition, nous pouvons voir que s'il existe un intervalle tel que $\Pi_a = 1$, qui implique que tous les autres intervalles ont une probabilité 0, alors l'hypothèse correspondant à l'intervalle Θ_a sera retenue, peu importe les valeurs des constantes.

Nous pouvons dès maintenant remarquer deux problèmes possibles avec cette fonction de perte. D'après la première condition, nous observons que nous choisissons toujours a_1 si k_{a_1} est suffisamment petit. En fait, en choisissant un k_{a_1} assez petit, nous pouvons nous assurer de choisir l'hypothèse H_{a_1} , peu importe les probabilités, même si l'hypothèse H_{a_1} a une probabilité *a posteriori* nulle. Toujours en regardant la première condition, nous remarquons que si nous faisons diminuer la largeur d'intervalle commune à tous les intervalles ε vers 0, et donc que les probabilités Π_a convergent vers 0, nous choisissons la décision à laquelle correspond la constante la plus petite. De ce fait, les comparaisons avec l'estimation ponctuelle sont douteuses car la règle de Bayes ne dépend plus de la densité *a posteriori* mais uniquement de la fonction de perte.

1.2.3. Fonction de perte L_4

Une autre façon de généraliser L_2 est d'assigner des pertes constantes spécifiques à la décision appropriée, au lieu de la décision choisie (c'est-à-dire L_3). Il va sans dire que la perte est nulle si la décision choisie correspond à la décision

appropriée. Voici donc notre quatrième fonction de perte,

$$L_4(\theta, d) = \begin{cases} 0 & \text{si } \theta \in \Theta_d; \\ k_e & \text{si } \theta \in \Theta_e, e \neq d, k_e > 0, \forall e \in \mathbb{Z}. \end{cases}$$

Il est important de noter que la perte (possible) dépend maintenant de la vraie valeur de θ , à travers la décision qui inclut cette valeur. Cette fonction de perte est aussi très flexible, étant donné que nous pouvons assigner les valeurs de k_e qui semblent appropriées.

Déterminons la décision de Bayes correspondant à cette fonction de perte. La perte espérée *a posteriori* est donnée par

$$\rho_4(\theta, a|x) = \mathbb{E}^{\pi(\theta|x)}[L_4(\theta, a)] = \left(\sum_{b \in \mathbb{Z}} k_b \Pi_b \right) - k_a \Pi_a$$

où $\Pi_b = \Pr\{\theta \in \Theta_b|x\}$.

Étant donné que la première partie de l'expression de droite ne dépend pas de la décision a , nous avons alors que $\rho_4(\theta, a|x)$ est minimal si et seulement si $k_a \Pi_a$ est maximal. Donc nous trouvons que la décision de Bayes est

$$\delta^\pi(x) = \arg \max_{a \in \mathbb{Z}} k_a \Pr\{\theta \in \Theta_a|x\}.$$

Les deux problèmes potentiels de la fonction de perte L_3 , soit le fait que l'on peut compenser une probabilité $\Pi_a = 0$ par une constante suffisamment petite et la tendance à choisir la décision correspondant à la plus petite constante lorsque $\varepsilon \rightarrow 0$, ne sont pas présentes pour L_4 . De plus, si un intervalle a une probabilité de 1, ce qui implique que tous les autres ont une probabilité de 0, cet intervalle sera toujours choisi, peu importe les constantes, en autant que $k_a > 0, \forall a \in \mathbb{Z}$.

1.3. FONCTIONS DE PERTE À VALEURS CONTINUES

1.3.1. Fonction de perte quadratique (L_5)

Notre cinquième fonction de perte devrait être familière pour les statisticiens, puisqu'elle est très similaire à l'erreur quadratique moyenne utilisée pour déterminer un estimateur pour un paramètre quelconque. Toutefois, elle doit être modifiée pour accommoder le fait que nous travaillons avec des intervalles et non pas avec des valeurs ponctuelles. L'équation de cette fonction de perte est donc

$$L_5(\theta, d) = \int_{\Theta_d} (t - \theta)^2 dt.$$

En développant cette formule, nous obtenons des résultats intéressants. Notons que $\Theta_d = (\nu_d - \frac{\varepsilon}{2}; \nu_d + \frac{\varepsilon}{2}]$ où ν_d est le centre de l'intervalle Θ_d et ε est la largeur commune à tous les intervalles. Ainsi,

$$\begin{aligned} L_5(\theta, d) &= \int_{\Theta_d} (t - \theta)^2 dt \\ &= \frac{1}{3} \left\{ \left(\nu_d + \frac{\varepsilon}{2} - \theta \right)^3 - \left(\nu_d - \frac{\varepsilon}{2} - \theta \right)^3 \right\} \\ &= \frac{1}{3} \left\{ \left[\left(\nu_d + \frac{\varepsilon}{2} \right)^3 - \left(\nu_d - \frac{\varepsilon}{2} \right)^3 \right] + 3\theta \left[\left(\nu_d + \frac{\varepsilon}{2} \right)^2 - \left(\nu_d - \frac{\varepsilon}{2} \right)^2 \right] \right. \\ &\quad \left. + 3\theta^2 \left[\left(\nu_d + \frac{\varepsilon}{2} \right) - \left(\nu_d - \frac{\varepsilon}{2} \right) \right] \right\} \\ &= \varepsilon \left\{ \nu_d^2 - 2\theta\nu_d + \theta^2 + \left(\frac{\varepsilon}{2} \right)^2 \right\} \\ &= \varepsilon \left\{ (\theta - \nu_d)^2 + \left(\frac{\varepsilon}{2} \right)^2 \right\} \\ &= \varepsilon \left\{ (\theta - \mathbb{E}^{\pi(\theta|x)}[\theta])^2 + (\theta - \mathbb{E}^{\pi(\theta|x)}[\theta]) (\mathbb{E}^{\pi(\theta|x)}[\theta] - \nu_d) \right. \\ &\quad \left. + (\mathbb{E}^{\pi(\theta|x)}[\theta] - \nu_d)^2 + \left(\frac{\varepsilon}{2} \right)^2 \right\}. \end{aligned}$$

Pour trouver la décision de Bayes pour cette fonction de perte, il nous faut la perte espérée *a posteriori*. Donc,

$$\begin{aligned}\rho_5(\theta, a|x) &= \mathbb{E}^{\pi(\theta|x)}[L_5(\theta, a)] \\ &= \varepsilon \left\{ \text{Var}^{\pi(\theta|x)}(\theta) + (\mathbb{E}^{\pi(\theta|x)}[\theta] - \nu_a)^2 + \left(\frac{\varepsilon}{2}\right)^2 \right\}\end{aligned}$$

étant donné que $\mathbb{E}^{\pi(\theta|x)}[\theta - \mathbb{E}^{\pi(\theta|x)}[\theta]] = 0$. Nous voyons que le seul terme qui dépend de la décision est le deuxième terme, que nous voulons donc minimiser. Puisque c'est une expression quadratique, nous avons que

$$\delta^\pi(x) = \arg \min_{a \in \mathbb{Z}} (\nu_a - \mathbb{E}^{\pi(\theta|x)}[\theta])^2 = \arg \min_{a \in \mathbb{Z}} |\nu_a - \mathbb{E}^{\pi(\theta|x)}[\theta]|.$$

Nous choisissons donc l'intervalle dont le centre est le plus près de l'espérance *a posteriori*.

1.3.2. Fonction de perte LINEX (intégrée) (L_6)

Présentée pour la première fois par Varian dans un article en 1975, cette dernière fonction de perte est à plusieurs égards une des plus intéressantes des six que nous présentons. Notons que LINEX tient pour LINéaire - EXponentielle. Nous nous basons sur l'article de Zellner (1986) pour les caractéristiques de cette fonction de perte.

En premier lieu, sa forme est la plus complexe et le développement de la perte espérée *a posteriori* (afin de l'optimiser) aura droit à son propre théorème. Le résultat a une forme élégante et « relativement » simple car elle dépend de la fonction génératrice de moments. Deuxièmement, sa forme (graphique) est approximativement linéaire d'un côté et exponentielle de l'autre. Ceci nous permet de pénaliser les erreurs d'estimation de façon asymétrique, ce qui peut se révéler très utile dans certains domaines d'application, dont l'actuariat.

L'équation de cette fonction de perte est

$$L_6(\theta, d) = \int_{\Theta_d} \{e^{\alpha(t-\theta)} - \alpha(t-\theta) - 1\} dt.$$

À l'intérieur des accolades, nous retrouvons une forme simplifiée de la fonction de perte LINEX (voir Varian(1975) ou Zellner(1986)). Le paramètre α sert à

contrôler la forme et l'« orientation » de la LINEX. Si $\alpha > 0$, le côté « exponentiel » est du côté de la surestimation de θ ($t > \theta$) et le côté « linéaire » est à l'opposé ($t < \theta$). Si $\alpha < 0$, c'est l'inverse. Pour des valeurs de $|\alpha|$ très petites, la LINEX ressemble à une fonction de perte quadratique (qui s'aplatit lorsque $|\alpha|$ diminue). Lorsque $|\alpha|$ est grand, on distingue clairement le côté exponentiel du côté linéaire. Ces observations sont illustrées dans la figure suivante.

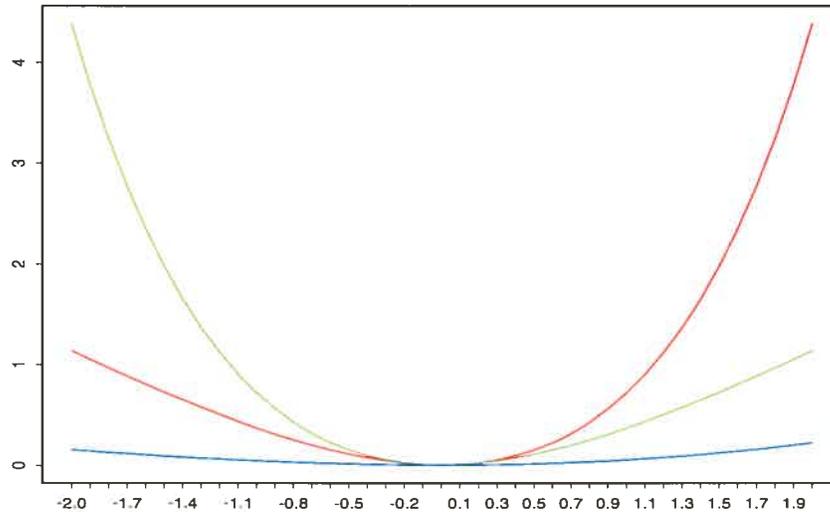


FIGURE 1.3.1. Illustration de la perte (pour la LINEX non intégrée) selon la valeur de $(d - \theta)$, pour 3 valeurs de α : $\alpha = 1$ (en rouge), $\alpha = -1$ (en vert) et $\alpha = 0.3$ (en bleu).

Théorème 1.3.1. *Si nous définissons $\Theta_a = (\nu_a - \frac{\varepsilon}{2}, \nu_a + \frac{\varepsilon}{2}]$, où ν_a est le centre et ε la largeur de l'intervalle, alors sous la fonction de perte*

$$L_G(\theta, d) = \int_{\Theta_a} \{e^{\alpha(t-\theta)} - \alpha(t-\theta) - 1\} dt.$$

l'intervalle ayant la perte espérée a posteriori minimale, noté $\Theta_{\delta\pi}$, sera celui tel que

$$\nu_{\delta\pi} \in \left(\frac{-1}{\alpha} \log(\mathbb{E}^{\pi(\theta|x)}[e^{-\alpha\theta}]) - \frac{3\varepsilon}{2} - \frac{2}{\alpha} \log\left(\frac{1 - e^{-\alpha\varepsilon}}{\alpha\varepsilon}\right) ; \right. \\ \left. \frac{-1}{\alpha} \log(\mathbb{E}^{\pi(\theta|x)}[e^{-\alpha\theta}]) - \frac{\varepsilon}{2} - \frac{2}{\alpha} \log\left(\frac{1 - e^{-\alpha\varepsilon}}{\alpha\varepsilon}\right) \right).$$

en supposant que $\alpha \neq 0$ et $\mathbb{E}^{\pi(\theta|x)}[e^{-\alpha\theta}]$ existe.

Remarque 1.3.1. Si $\alpha = 0$, nous avons alors que $L(\theta, d) \equiv 0$, une fonction de perte peu intéressante s'il en est une.

Afin d'alléger la notation et de faciliter la comparaison avec l'estimation ponctuelle, nous déterminons la forme de l'estimateur ponctuel avant de procéder à la preuve du théorème (1.3.1).

Proposition 1.3.1. L'estimateur ponctuel de la fonction de perte LINEX (non intégrée) est

$$\hat{\theta}_B = \frac{-1}{\alpha} \log (\mathbb{E}^{\pi(\theta|x)}[e^{-\alpha\theta}]).$$

Démonstration (proposition 1.3.1) :

La fonction de perte est

$$L(\theta, d) = e^{\alpha(d-\theta)} - \alpha(d - \theta) - 1.$$

Montrons que $\hat{\theta}_B$ minimise la perte espérée *a posteriori*, qui est donnée par

$$\rho(\theta, t|x) = e^{\alpha t} \mathbb{E}^{\pi(\theta|x)}[e^{-\alpha\theta}] - \alpha t + \alpha \mathbb{E}^{\pi(\theta|x)}[\theta] - 1.$$

Ainsi, pour trouver notre estimateur, nous devons résoudre

$$\begin{aligned} \frac{d}{dt} \rho(\theta, t|x) &= \alpha e^{\alpha t} \mathbb{E}^{\pi(\theta|x)}[e^{-\alpha\theta}] - \alpha = 0 \\ \Leftrightarrow e^{-\alpha t} &= \mathbb{E}^{\pi(\theta|x)}[e^{-\alpha\theta}] \\ \Leftrightarrow t &= \frac{-1}{\alpha} \log (\mathbb{E}^{\pi(\theta|x)}[e^{-\alpha\theta}]) (= \hat{\theta}_B). \end{aligned}$$

Nous avons bien un minimum puisque

$$\frac{d^2}{dt^2} \rho(\theta, t|x) = \alpha^2 e^{\alpha t} \mathbb{E}^{\pi(\theta|x)}[e^{-\alpha\theta}] > 0, \quad \forall t$$

en supposant que $\alpha \neq 0$.

Démonstration (théorème 1.3.1) :

Nous commençons par trouver une expression pour la perte sous la fonction de perte LINEX, celle-ci étant

$$\begin{aligned}
L_6(\theta, d) &= \int_{\Theta_d} [e^{\alpha(t-\theta)} - \alpha(t-\theta) - 1] dt \\
&= e^{-\alpha\theta} \left(\frac{e^{\alpha(\nu_d + \frac{\varepsilon}{2})} - e^{\alpha(\nu_d - \frac{\varepsilon}{2})}}{\alpha} \right) - \frac{\alpha}{2} \left[(\nu_d + \frac{\varepsilon}{2} - \theta)^2 \right. \\
&\quad \left. - (\nu_d - \frac{\varepsilon}{2} - \theta)^2 \right] - \varepsilon \\
&= \varepsilon \left[\left(\frac{e^{\frac{\alpha\varepsilon}{2}} - e^{-\frac{\alpha\varepsilon}{2}}}{\alpha\varepsilon} \right) e^{\alpha(\nu_d - \theta)} - \alpha(\nu_d - \theta) - 1 \right].
\end{aligned}$$

Nous retrouvons donc la fonction de perte LINEX (non intégrée) évaluée au centre de l'intervalle avec une légère modification au terme exponentiel (il est multiplié par $\frac{\sinh(\frac{\alpha\varepsilon}{2})}{(\frac{\alpha\varepsilon}{2})}$), le tout multiplié par la largeur de l'intervalle ε .

Passons maintenant à la perte espérée *a posteriori* :

$$\begin{aligned}
\rho_6(\theta, a|x) &= \mathbb{E}^{\pi(\theta|x)}[L_6(\theta, a)] \\
&= \varepsilon \left[\left(\frac{e^{\frac{\alpha\varepsilon}{2}} - e^{-\frac{\alpha\varepsilon}{2}}}{\alpha\varepsilon} \right) e^{\alpha\nu_a} \mathbb{E}^{\pi(\theta|x)}[e^{-\alpha\theta}] - \alpha\nu_a + \alpha\mathbb{E}^{\pi(\theta|x)}[\theta] - 1 \right] \\
&= \varepsilon \left[\left(\frac{e^{\frac{\alpha\varepsilon}{2}} - e^{-\frac{\alpha\varepsilon}{2}}}{\alpha\varepsilon} \right) e^{\alpha(\nu_a - \hat{\theta}_B)} - \alpha(\nu_a - \hat{\theta}_B) - 1 \right] \\
&\quad + \varepsilon \left[\alpha(\mathbb{E}^{\pi(\theta|x)}[\theta] - \hat{\theta}_B) \right]
\end{aligned}$$

où $\hat{\theta}_B = \frac{-1}{\alpha} \log(\mathbb{E}^{\pi(\theta|x)}[e^{-\alpha\theta}])$ (voir proposition 1.3.1). Seuls les deux premiers termes de l'expression finale nous intéressent, puisque la deuxième partie est constante peu importe la décision a .

Bien que les ν_a soient fixes pour chaque situation particulière pour tous les intervalles Θ_a , nous pouvons tout de même tenter de caractériser le comportement de la fonction de perte espérée *a posteriori*, surtout autour de son minimum, si celui existe. Pour ce faire, nous allons supposer que ν_a est une variable continue et dériver la perte espérée *a posteriori* par rapport à celle-ci. Ceci nous permettra

de trouver pour quelle valeur de centre (d'intervalle) la perte espérée *a posteriori* est minimale. Nous noterons cette valeur par $\hat{\nu}$. Notons que $\hat{\nu}$ est le minimum sur toutes les valeurs de \mathbb{R} et n'est pas limité aux valeurs des centres d'intervalle. Ainsi, nous obtenons

$$\begin{aligned} \frac{\partial}{\partial \nu_a} \rho_6(\theta, a|x) &= \alpha \varepsilon \left[\left(\frac{e^{\frac{\alpha \varepsilon}{2}} - e^{-\frac{\alpha \varepsilon}{2}}}{\alpha \varepsilon} \right) e^{\alpha(\hat{\nu} - \hat{\theta}_B)} - 1 \right] = 0 \\ \Leftrightarrow e^{\alpha(\hat{\nu} - \hat{\theta}_B)} &= \left(\frac{e^{\frac{\alpha \varepsilon}{2}} - e^{-\frac{\alpha \varepsilon}{2}}}{\alpha \varepsilon} \right)^{-1} \\ \Leftrightarrow \hat{\nu} &= \hat{\theta}_B - \frac{1}{\alpha} \log \left(\frac{e^{\frac{\alpha \varepsilon}{2}} - e^{-\frac{\alpha \varepsilon}{2}}}{\alpha \varepsilon} \right). \end{aligned} \quad (1.3.1)$$

Remarque 1.3.2.

$$\begin{aligned} \left(\frac{e^{\frac{\alpha \varepsilon}{2}} - e^{-\frac{\alpha \varepsilon}{2}}}{\alpha \varepsilon} \right) &= 1 + \frac{(\alpha \varepsilon)^2}{24} + \frac{(\alpha \varepsilon)^2}{1920} + \dots \\ &= 1 + O((\alpha \varepsilon)^2). \end{aligned}$$

On voit que le résultat converge rapidement vers 1 lorsque $\varepsilon \rightarrow 0$, ce qui implique que $\hat{\nu} \rightarrow \hat{\theta}_B - \frac{1}{\alpha} O((\alpha \varepsilon)^2)$.

Ainsi, en utilisant l'équation (1.3.1) nous avons

$$\begin{aligned} \frac{\partial}{\partial \nu_a} \rho_6(\theta, a|x) &= \alpha \varepsilon \left[\left(\frac{e^{\frac{\alpha \varepsilon}{2}} - e^{-\frac{\alpha \varepsilon}{2}}}{\alpha \varepsilon} \right) e^{\alpha(\nu_a - \hat{\theta}_B)} - 1 \right] \\ &= \alpha \varepsilon [e^{\alpha(\nu_a - \hat{\nu})} - 1]. \end{aligned}$$

Donc, nous avons que

$$\frac{\partial}{\partial \nu_a} \rho_6(\theta, a|x) \begin{cases} > 0 \quad \forall \nu_a > \hat{\nu}; \\ < 0 \quad \forall \nu_a < \hat{\nu}; \\ = 0, \nu_a = \hat{\nu}. \end{cases}$$

À partir des résultats précédents, nous pouvons chercher une expression pour spécifier l'intervalle ayant la perte espérée *a posteriori* minimale.

Sauf coïncidence, aucun intervalle n'a pour centre exactement $\hat{\nu}$, alors d'après ce que nous venons de voir à propos de la « dérivée » de la perte espérée *a posteriori*, les deux seuls intervalles qui peuvent avoir la perte espérée *a posteriori* minimale sont ceux dont les centres sont immédiatement à droite (que nous noterons ν_d pour Θ_d) et à gauche (que nous noterons ν_g pour Θ_g) de $\hat{\nu}$.

Rappelons que nous assumons que les intervalles sont de largeur constante, notée ε . Ainsi, notons

$$\varepsilon_d = \nu_d - \hat{\nu} > 0,$$

$$\varepsilon_g = \nu_g - \hat{\nu} = \varepsilon_d - \varepsilon < 0,$$

car $\nu_d - \nu_g = \varepsilon$ et $\nu_g < \hat{\nu} < \nu_d$. Cherchons maintenant d^* (notons que d^* n'est pas nécessairement un entier) tel que

$$\rho_6(\theta, d^* | x) = \rho_6(\theta, d^* - 1 | x)$$

avec $\nu_{d^*} = \nu_{\lfloor d^* \rfloor} + (d^* - \lfloor d^* \rfloor)\varepsilon$ et $\varepsilon_{d^*} = \nu_{d^*} - \hat{\nu}$ (notons que $\varepsilon_{d^*} < \varepsilon$).

Nous pouvons considérer ν_{d^*} comme le point auquel nous sommes neutre entre l'intervalle de gauche et celui de droite (les deux ayant la même perte espérée *a posteriori*). Si l'intervalle de droite a son centre $\nu_d < \nu_{d^*}$ (rappelons que $\hat{\nu} < \nu_d$), la perte espérée *a posteriori* de l'intervalle de droite est inférieure à celle de l'intervalle de gauche (voir les comportements de la perte espérée *a posteriori* et sa « dérivée »). Sinon, nous choisirons l'intervalle de gauche. En utilisant encore une fois l'équation (1.3.1) pour simplifier le terme devant l'exponentielle, nous avons donc

$$\begin{aligned}
\rho_6(\theta, d^*|x) &= \rho_6(\theta, d^* - 1|x) \\
&\Leftrightarrow \left(\frac{e^{\frac{\alpha\varepsilon}{2}} - e^{-\frac{\alpha\varepsilon}{2}}}{\alpha\varepsilon} \right) e^{\alpha(\nu_{d^*} - \hat{\nu})} e^{\alpha(\hat{\nu} - \hat{\theta}_B)} - \alpha(\nu_{d^*} - \hat{\nu}) \\
&= \left(\frac{e^{\frac{\alpha\varepsilon}{2}} - e^{-\frac{\alpha\varepsilon}{2}}}{\alpha\varepsilon} \right) e^{\alpha(\nu_{d^*} - \varepsilon - \hat{\nu})} e^{\alpha(\hat{\nu} - \hat{\theta}_B)} - \alpha(\nu_{d^*} - \varepsilon - \hat{\nu}) \\
&\Leftrightarrow e^{\alpha\varepsilon_d^*} - \alpha\varepsilon_d^* = e^{\alpha(\varepsilon_d^* - \varepsilon)} - \alpha(\varepsilon_d^* - \varepsilon) \\
&\Leftrightarrow (1 - e^{-\alpha\varepsilon}) e^{\varepsilon_d^*} = \alpha\varepsilon \\
&\Leftrightarrow \varepsilon_d^* = \frac{1}{\alpha} [\log(\alpha\varepsilon) - \log(1 - e^{-\alpha\varepsilon})] \\
&\Leftrightarrow \nu_{d^*} = \hat{\nu} + \frac{1}{\alpha} [\log(\alpha\varepsilon) - \log(1 - e^{-\alpha\varepsilon})] \\
&= \left[\frac{-1}{\alpha} \log(\mathbb{E}^{\pi(\theta|x)}[e^{-\alpha\theta}]) - \frac{1}{\alpha} \log\left(e^{\frac{\alpha\varepsilon}{2}} \left(\frac{1 - e^{-\alpha\varepsilon}}{\alpha\varepsilon} \right) \right) \right] \\
&\quad - \frac{1}{\alpha} \log\left(\frac{1 - e^{-\alpha\varepsilon}}{\alpha\varepsilon} \right) \\
&= \frac{-1}{\alpha} \log(\mathbb{E}^{\pi(\theta|x)}[e^{-\alpha\theta}]) - \frac{\varepsilon}{2} - \frac{2}{\alpha} \log\left(\frac{1 - e^{-\alpha\varepsilon}}{\alpha\varepsilon} \right).
\end{aligned}$$

Puisque nous choisissons l'intervalle de droite si $\nu_d < \nu_{d^*}$ et l'intervalle de gauche si $\nu_g > \nu_{d^*} - \varepsilon$ ($\Leftrightarrow \nu_d > \nu_{d^*}$), nous pouvons conclure que l'intervalle choisi $\Theta_{\delta^\pi(x)}$ sera celui tel que

$$\begin{aligned}
\nu_{\delta^\pi(x)} \in & \left(\frac{-1}{\alpha} \log(\mathbb{E}^{\pi(\theta|x)}[e^{-\alpha\theta}]) - \frac{3\varepsilon}{2} - \frac{2}{\alpha} \log\left(\frac{1 - e^{-\alpha\varepsilon}}{\alpha\varepsilon} \right) ; \right. \\
& \left. \frac{-1}{\alpha} \log(\mathbb{E}^{\pi(\theta|x)}[e^{-\alpha\theta}]) - \frac{\varepsilon}{2} - \frac{2}{\alpha} \log\left(\frac{1 - e^{-\alpha\varepsilon}}{\alpha\varepsilon} \right) \right].
\end{aligned}$$

Nous avons arbitrairement choisi de fermer à droite pour éviter la situation où les centres des intervalles de gauche et de droite tombent respectivement sur la borne de gauche et la borne de droite. Ayant fermé du côté droit, nous choisirons alors l'intervalle de droite même si les deux intervalles ont la même perte espérée *a posteriori*.

1.4. CONCLUSION

Nous avons commencé par présenter les notions statistiques de base dont nous avons besoin tout au long de ce mémoire, soit les fonctions de perte, les tests d'hypothèses multiples, l'estimation ponctuelle et un minimum de statistique bayésienne. Nous avons aussi présenté les caractéristiques essentielles de notre partition de l'espace-paramètre.

Ensuite, nous avons présenté tour à tour six fonctions de perte (dont deux reviennent au même, soit L_1 et L_2), divisées selon qu'elles spécifient des pertes prédéfinies (ou 0 si l'on ne se trompe pas), c'est-à-dire L_1 à L_4 ou des pertes pouvant prendre n'importe quelle valeur réelle positive, soit L_5 et L_6 .

Pour chacune de ces fonctions de perte, nous avons déterminé quelle sera la décision de Bayes qui entraîne la perte espérée *a posteriori* minimale, et donc quel intervalle sera choisi, parmi notre partition.

Le chapitre qui suit appliquera ces résultats à des situations un peu plus spécifiques, soit des types de distribution caractérisées par le nombre de modes et la symétrie de la fonction de densité.

Chapitre 2

DISTRIBUTIONS ET RÉSULTATS SPÉCIFIQUES

Nous appliquons maintenant les résultats obtenus au premier chapitre à trois types de distribution : unimodale symétrique (US), unimodale asymétrique (UA) et bimodale (BI). Nous tentons de pousser l'analyse le plus loin possible dans chaque cas, mais la généralité de ces catégories nous limite grandement et nous ne sommes pas intéressés à dresser une liste interminable de distributions et de leurs résultats particuliers. Des résultats plus spécifiques devraient pouvoir être obtenus sans grande difficulté pour la plupart des distributions les plus populaires.

Pour terminer le chapitre, nous présentons un théorème très général qui limite le nombre d'intervalles à « examiner » afin de trouver celui ayant la probabilité *a posteriori* maximale et ce, pour toute distribution ayant une densité continue et un nombre de modes fini. En fait, les résultats des trois types de distribution correspondants à la fonction de perte L_1 (et donc aussi à L_2) s'obtiennent à partir de ce théorème général, puisqu'ils sont des cas particuliers à 1 et 2 modes. Dans tous les cas, nous assumons que les distributions sont continues et les largeurs des intervalles sont constantes et fixées à une valeur ε .

2.1. DISTRIBUTION UNIMODALE SYMÉTRIQUE

2.1.1. Caractéristiques

Par distribution unimodale symétrique, nous entendons une distribution ayant un mode unique (qui peut être un plateau, auquel cas le mode est considéré comme étant le milieu) et dont la fonction de densité est symétrique autour de ce mode.

Ceci a pour implication que le mode, la médiane et la moyenne (l'espérance), si elle existe, prennent tous la même valeur, que nous notons θ_0 . Des exemples de distributions unimodales symétriques sont la loi normale, la loi t de Student et la double exponentielle.

Précisons que les intervalles ne sont pas nécessairement placés de telle façon que le mode soit au centre d'un intervalle.

2.1.2. Résultats

2.1.2.1. Fonctions de perte 0-1 et 0-k (L_1 et L_2)

Rappelons que

$$L_1(\theta, d) = \begin{cases} 0 & \text{si } \theta \in \Theta_d; \\ 1 & \text{sinon,} \end{cases}$$

et que la décision de Bayes correspondant à L_1 est $\delta^\pi(x) = \arg \max_{a \in \mathbb{Z}} \Pr\{\theta \in \Theta_a | x\}$.

Théorème 2.1.1. *La forme de la distribution (unimodale symétrique) implique que son mode $\theta_0 \in \Theta_{\delta^\pi(x)} = (l_{\delta^\pi(x)}, u_{\delta^\pi(x)}]$.*

Démonstration (théorème 2.1.1) :

Dans le but d'alléger la notation, nous utilisons $\delta^\pi := \delta^\pi(x)$.

Afin de démontrer ce résultat, choisissons un intervalle quelconque $\Theta_b = (l_b, u_b]$ et assumons sans perdre de généralité que $l_b \geq u_{\delta^\pi}$ (donc que Θ_b « est à droite »

de $\Theta_{\delta\pi}$). Puisque la densité décroît à mesure que l'on s'éloigne de θ_0 , nous savons que $\forall y \in \Theta_b, \pi(y|x) \leq \pi(l_b|x) \leq \pi(u_{\delta\pi}|x)$.

Ici, deux possibilités s'offrent à nous, la première et la plus simple étant que $\pi(l_{\delta\pi}|x) \geq \pi(u_{\delta\pi}|x)$, ce qui implique que $\theta_0 - l_{\delta\pi} \leq u_{\delta\pi} - \theta_0 \Leftrightarrow l_{\delta\pi} + u_{\delta\pi} \geq 2\theta_0$, puisque la densité est symétrique autour de θ_0 . Notons en passant que $\frac{1}{2}(l_{\delta\pi} + u_{\delta\pi})$ est le centre de l'intervalle $\Theta_{\delta\pi} = (l_{\delta\pi}, u_{\delta\pi}]$. Nous procédons alors par

$$\begin{aligned} \Pr\{\theta \in \Theta_{\delta\pi}|x\} &\geq \varepsilon\pi(u_{\delta\pi}|x) \\ &\geq \varepsilon\pi(l_b|x) \\ &\geq \Pr\{\theta \in \Theta_b|x\}. \end{aligned}$$

La deuxième possibilité est que $\pi(l_{\delta\pi}|x) < \pi(u_{\delta\pi}|x)$ (et donc que $l_{\delta\pi} + u_{\delta\pi} < 2\theta_0$). De nouveau, il y a deux cas possibles. Si $\pi(l_{\delta\pi}|x) \geq \pi(l_b|x)$, donc $l_b - \theta_0 \geq \theta_0 - l_{\delta\pi}$ ($\Leftrightarrow l_{\delta\pi} + l_b \geq 2\theta_0$), alors nous procédons similairement à la possibilité précédente, c'est-à-dire

$$\begin{aligned} \Pr\{\theta \in \Theta_{\delta\pi}|x\} &\geq \varepsilon\pi(l_{\delta\pi}|x) \\ &\geq \varepsilon\pi(l_b|x) \\ &\geq \Pr\{\theta \in \Theta_b|x\}. \end{aligned}$$

Sinon, $\pi(l_{\delta\pi}|x) < \pi(l_b|x)$. Ceci implique que $l_b + l_{\delta\pi} < 2\theta_0 \Leftrightarrow l_b = u_{\delta\pi}$. Ceci vient du fait que $\theta_0 \in \Theta_{\delta\pi}$ et donc $l_{\delta\pi} > \theta_0 - \varepsilon$. Dans ce cas, la preuve devient

$$\begin{aligned} \Pr\{\theta \in \Theta_{\delta\pi}|x\} &= \int_{l_{\delta\pi}}^{u_{\delta\pi}} \pi(\theta|x)d\theta \\ &= \int_{l_{\delta\pi}}^{\theta_0 - (u_{\delta\pi} - \theta_0)} \pi(\theta|x)d\theta + \int_{\theta_0 - (u_{\delta\pi} - \theta_0)}^{u_{\delta\pi}} \pi(\theta|x)d\theta. \end{aligned}$$

Notons maintenant que l'intervalle $(u_{\delta\pi} = l_b, 2\theta_0 - l_{\delta\pi})$ est symétrique à l'intervalle $(l_{\delta\pi}, \theta_0 - (u_{\delta\pi} - \theta_0))$ par rapport au mode θ_0 . Ceci implique que les intégrales de la fonction de densité sur ces deux intervalles donnent des résultats égaux,

puisque la densité est symétrique autour du mode. Nous obtenons alors

$$\begin{aligned}
& \int_{l_{\delta\pi}}^{\theta_0 - (u_{\delta\pi} - \theta_0)} \pi(\theta|x) d\theta + \int_{\theta_0 - (u_{\delta\pi} - \theta_0)}^{u_{\delta\pi}} \pi(\theta|x) d\theta \\
&= \int_{u_{\delta\pi} = l_b}^{2\theta_0 - l_{\delta\pi}} \pi(\theta|x) d\theta + \int_{2\theta_0 - u_{\delta\pi}}^{u_{\delta\pi}} \pi(\theta|x) d\theta \\
&\geq \int_{l_b}^{2\theta_0 - l_{\delta\pi}} \pi(\theta|x) d\theta + 2(u_{\delta\pi} - \theta_0) \pi(u_{\delta\pi}|x) \\
&\geq \int_{l_b}^{2\theta_0 - l_{\delta\pi}} \pi(\theta|x) d\theta + \int_{2\theta_0 - l_{\delta\pi}}^{u_{\delta\pi} + \varepsilon = u_b} \pi(\theta|x) d\theta \\
&= \Pr\{\theta \in \Theta_b|x\}.
\end{aligned}$$

Nous pouvons donc conclure que $\theta_0 \in \Theta_{\delta\pi}$ car si Θ_b « est à gauche » de $\Theta_{\delta\pi}$, la démonstration est pratiquement la même étant donné que la distribution est symétrique.

2.1.2.2. Fonction de perte L_3

Rappelons que

$$L_3(\theta, d) = \begin{cases} 0 & \text{si } \theta \in \Theta_d; \\ k_d & \text{sinon} \end{cases},$$

et que la décision de Bayes pour L_3 est $\delta^\pi(x) = \arg \min_{a \in \mathbb{Z}} k_a (1 - \Pr\{\theta \in \Theta_a|x\})$.

Étant donné qu'il est impossible de déterminer de façon générale l'hypothèse retenue à cause de l'influence déterminante des constantes k_a , nous n'étudions que quelques conséquences de ce choix de fonction de perte.

Nous commençons par comparer le résultat pour L_3 avec celui pour L_1 et L_2 . Nous savons déjà que L_1 et L_2 nous font choisir $H_m : \theta \in \Theta_m$ tel que $\theta_0 \in \Theta_m$. Par la première condition trouvée à la section 1.2.2, H_m est rejetée si $\exists b \neq m$ tel que

$$k_b < k_m \frac{1 - \Pr\{\theta \in \Theta_m|x\}}{1 - \Pr\{\theta \in \Theta_b|x\}}. \quad (2.1.1)$$

Cette borne supérieure sera maximale pour la plus grande $\Pr\{\theta \in \Theta_b|x\}$, en exceptant l'intervalle retenu par L_1 et L_2 . Celle-ci se trouve pour un des deux intervalles $\Theta_{m\pm 1}$. En fait, si $\pi(l_m|x) < \pi(u_m|x)$, ce sera $\Theta_{m+1} = (l_m + \varepsilon, u_m + \varepsilon]$. Si nous avons $\pi(l_m|x) > \pi(u_m|x)$, ce sera $\Theta_{m-1} = (l_m - \varepsilon, u_m - \varepsilon]$. Si $\pi(l_m|x) = \pi(u_m|x)$, $\Pr\{\theta \in \Theta_{m-1}|x\} = \Pr\{\theta \in \Theta_{m+1}|x\}$ par symétrie. Donc, nous pouvons conclure que s'il n'existe pas d'entier $b \neq m$ tel que

$$k_b < k_m \frac{1 - \Pr\{\theta \in \Theta_m|x\}}{1 - \max\{\Pr\{\theta \in \Theta_{m-1}|x\}, \Pr\{\theta \in \Theta_{m+1}|x\}\}}$$

alors $\delta^\pi = m$ pour L_3 , qui s'accorde dans ce cas avec L_1 et L_2 .

2.1.2.3. Fonction de perte L_4

Rappelons que

$$L_4(\theta, d) = \begin{cases} 0 & \text{si } \theta \in \Theta_d; \\ k_b & \text{si } \theta \in \Theta_b, b \neq d \end{cases},$$

et que la décision de Bayes est donnée par $\delta^\pi(x) = \arg \max_{a \in \mathbb{Z}} k_a \Pr\{\theta \in \Theta_a|x\}$.

Comme pour L_3 , l'importance des constantes k_a est trop grande pour pouvoir déterminer l'hypothèse qui sera retenue. Nous pouvons tout de même tirer quelques conclusions générales.

Comparons le résultat dû à L_4 à celui de L_1 et L_2 . Soit $H_m : \theta \in \Theta_m$ tel que $\theta_0 \in \Theta_m$.

Puisque H_m est rejetée si $\exists b \neq m$ tel que

$$k_b > k_m \frac{\Pr\{\theta \in \Theta_m|x\}}{\Pr\{\theta \in \Theta_b|x\}}, \quad (2.1.2)$$

nous pouvons conclure, de façon similaire à L_3 précédemment, que nous choisissons H_m s'il n'existe pas d'entier $b \neq m$ tel que

$$k_b > k_m \frac{\Pr\{\theta \in \Theta_m|x\}}{\max\{\Pr\{\theta \in \Theta_{m-1}|x\}, \Pr\{\theta \in \Theta_{m+1}|x\}\}}.$$

Dans cette éventualité, L_4 entraîne le choix de la même hypothèse que L_1 et L_2 .

2.1.2.4. Fonction de perte quadratique (L_5)

Rappelons que

$$L_5(\theta, d) = \int_{\Theta_a} (t - \theta)^2 dt$$

et la décision de Bayes correspondant à cette fonction de perte est $\delta^\pi(x) = \arg \min_{a \in \mathcal{Z}} (\nu_a - \mathbb{E}^{\pi(\theta|x)}[\theta])^2$, où ν_a est le centre de l'intervalle $\Theta_a = (l_a, u_a] = (\nu_a - \frac{\varepsilon}{2}, \nu_a + \frac{\varepsilon}{2}]$.

Dans le cas d'une distribution unimodale symétrique, nous savons que $\mathbb{E}^{\pi(\theta|x)}[\theta] = \theta_0$, si $\mathbb{E}^{\pi(\theta|x)}[\theta]$ existe.

Donc, $\theta_0 \in \Theta_{\delta^\pi}$ et L_5 s'accorde avec L_1 et L_2 dans tous les cas de distribution unimodale symétrique (encore une fois, si la moyenne existe).

2.1.2.5. Fonction de perte LINEX (L_6)

Rappelons que

$$L_6(\theta, d) = \int_{\Theta_a} [e^{\alpha(t-\theta)} - \alpha(t-\theta) - 1] dt$$

et la décision de Bayes est telle que

$$\nu_{\delta^\pi(x)} \in \left(\frac{-1}{\alpha} \log(\mathbb{E}^{\pi(\theta|x)}[e^{-\alpha\theta}]) - \frac{3\varepsilon}{2} - \frac{2}{\alpha} \log\left(\frac{1 - e^{-\alpha\varepsilon}}{\alpha\varepsilon}\right) : \right. \\ \left. \frac{-1}{\alpha} \log(\mathbb{E}^{\pi(\theta|x)}[e^{-\alpha\theta}]) - \frac{\varepsilon}{2} - \frac{2}{\alpha} \log\left(\frac{1 - e^{-\alpha\varepsilon}}{\alpha\varepsilon}\right) \right]$$

où $\nu_{\delta^\pi(x)}$ est le centre de l'intervalle $\Theta_{\delta^\pi(x)} = (l_{\delta^\pi}, u_{\delta^\pi}] = (\nu_{\delta^\pi(x)} - \frac{\varepsilon}{2}, \nu_{\delta^\pi(x)} + \frac{\varepsilon}{2}]$.

Dû à la symétrie de la distribution, nous pouvons conclure que si $\alpha > 0$, alors $\hat{\theta}_B = \frac{-1}{\alpha} \log(\mathbb{E}^{\pi(\theta|x)}[e^{-\alpha\theta}]) < \theta_0$ et l'inverse lorsque $\alpha < 0$.

2.2. DISTRIBUTION UNIMODALE ASYMÉTRIQUE

2.2.1. Caractéristiques

Par distribution unimodale asymétrique, nous entendons une distribution ayant un mode unique (qui peut être un plateau) et dont la fonction de densité n'est pas symétrique autour de ce mode. Nous assumons que la fonction de densité est continue.

Contrairement à la distribution unimodale symétrique, le mode θ_0 , la médiane et la moyenne (l'espérance), si elle existe, ne prennent pas la même valeur. Des exemples de distribution unimodale asymétrique sont la loi gamma et la loi lognormale. Il est à noter que la fonction de perte L_6 n'est pas applicable à la loi lognormale puisque celle-ci n'a pas de fonction génératrice de moments. Notons aussi que la loi exponentielle ne respecte pas nos conditions sur la densité et donc que les résultats pour L_1 et L_2 ne s'y appliquent pas.

2.2.2. Résultats

2.2.2.1. Fonctions de perte 0 – 1 et 0 – k (L_1 et L_2)

Rappelons que $\delta^\pi(x) = \arg \max_{a \in \mathbb{Z}} \Pr\{\theta \in \Theta_a | x\}$.

Soit Θ_α tel que $\theta_0 \in \Theta_\alpha = (l_\alpha, u_\alpha]$. Nous montrons maintenant qu'il n'existe que deux intervalles qui peuvent être l'intervalle de probabilité *a posteriori* maximale.

Il existe trois cas possibles :

- $\pi(l_\alpha | x) = \pi(u_\alpha | x)$, on choisit alors Θ_α ($\delta^\pi = \alpha$);
- $\pi(l_\alpha | x) < \pi(u_\alpha | x)$, les deux possibilités sont $\Theta_\alpha = (l_\alpha, u_\alpha]$ et $\Theta_{\alpha+1} = (l_{\alpha+1}, u_{\alpha+1}] = (l_\alpha + \varepsilon, u_\alpha + \varepsilon]$;
- $\pi(l_\alpha | x) > \pi(u_\alpha | x)$, les deux possibilités sont $\Theta_\alpha = (l_\alpha, u_\alpha]$ et $\Theta_{\alpha-1} = (l_{\alpha-1}, u_{\alpha-1}] = (l_\alpha - \varepsilon, u_\alpha - \varepsilon]$.

Puisque la preuve dans le 3^e cas est pratiquement la même que pour le 2^e, nous l'omettrons.

Prouvons le 2^e cas. Pour alléger la notation, nous noterons $\beta_\alpha = \alpha + 1$.

Puisque la distribution est unimodale avec θ_0 comme mode unique, nous savons que

$$\pi(\theta_0|x) \geq \pi(s|x) \geq \pi(t|x)$$

lorsque $\theta_0 \leq s \leq t$ ou $t \leq s \leq \theta_0$.

Pour un intervalle quelconque Θ_a « à gauche » de Θ_α , c'est-à-dire $u_a \leq l_\alpha$, nous avons que

$$\begin{aligned} \Pr\{\theta \in \Theta_\alpha|x\} &= \int_{l_\alpha}^{u_\alpha} \pi(\theta|x)d\theta \\ &= \int_{l_\alpha}^{\theta_0} \pi(\theta|x)d\theta + \int_{\theta_0}^{u_\alpha} \pi(\theta|x)d\theta \\ &\geq (\theta_0 - l_\alpha)\pi(l_\alpha|x) + (u_\alpha - \theta_0)\pi(u_\alpha|x) \\ &> \varepsilon\pi(l_\alpha|x) \\ &\geq \varepsilon\pi(u_a|x) \\ &\geq \int_{l_a}^{u_a} \pi(\theta|x)d\theta \\ &= \Pr\{\theta \in \Theta_a|x\}. \end{aligned}$$

Pour un intervalle quelconque Θ_a « à droite » de Θ_{β_α} , c'est-à-dire $l_a \geq u_{\beta_\alpha}$, nous avons que

$$\begin{aligned} \Pr\{\theta \in \Theta_{\beta_\alpha} | x\} &= \int_{l_{\beta_\alpha}}^{u_{\beta_\alpha}} \pi(\theta | x) d\theta \\ &\geq \varepsilon \pi(u_{\beta_\alpha} | x) \\ &\geq \varepsilon \pi(l_a | x) \\ &\geq \int_{l_a}^{u_a} \pi(\theta | x) d\theta \\ &= \Pr\{\theta \in \Theta_a | x\}. \end{aligned}$$

Pour le premier cas, la preuve revient à la même chose que la première partie du deuxième cas, dans les deux directions, en changeant le ">" pour un "≥".

2.2.2.2. Fonction de perte L_3

Rappelons que $\delta^\pi(x) = \arg \min_{a \in \mathbb{Z}} k_a (1 - \Pr\{\theta \in \Theta_a | x\})$.

Comme pour la distribution unimodale symétrique, nous devons nous satisfaire de quelques résultats à cause de l'impossibilité de tirer des conclusions générales.

Notons par H_m l'hypothèse retenue par les fonctions de perte L_1 et L_2 . Nous pouvons alors affirmer que $\delta^\pi(x) = m$ pour L_3 si $\nexists b \neq m$ tel que

$$k_b < k_m \frac{1 - \Pr\{\theta \in \Theta_m | x\}}{1 - \max\{\Pr\{\theta \in \Theta_{m-1} | x\}, \Pr\{\theta \in \Theta_{m+1} | x\}\}}.$$

Il faut noter que nous n'avons pas nécessairement que $\theta_0 \in \Theta_m$.

2.2.2.3. Fonction de perte L_4

Rappelons que $\delta^\pi(x) = \arg \max_{a \in \mathbb{Z}} k_a \Pr\{\theta \in \Theta_a | x\}$.

Comme pour L_3 , nous nous contentons de comparer le résultat de L_4 avec celui de L_1 et L_2 .

Notons par H_m que l'hypothèse retenue par L_1 et L_2 . Nous avons alors que L_3 choisit H_m s'il n'existe pas d'entier $b \neq m$ tel que

$$k_b > k_m \frac{\Pr\{\theta \in \Theta_m | x\}}{\max\{\Pr\{\theta \in \Theta_{m-1} | x\}, \Pr\{\theta \in \Theta_{m+1} | x\}\}}.$$

Encore une fois, il se peut que nous n'ayons pas $\theta_0 \in \Theta_m$.

2.2.2.4. Fonction de perte quadratique (L_5)

Rappelons que $\delta^\pi(x) = \arg \min_{a \in \mathbb{Z}} (\nu_a - \mathbb{E}^{\pi(\theta|x)}[\theta])^2$.

Mais, puisque nous n'avons pas nécessairement $\theta_0 = \mathbb{E}^{\pi(\theta|x)}[\theta]$ et que $\mathbb{E}^{\pi(\theta|x)}[\theta] \in \Theta_{\delta^\pi}$, il se peut que nous n'ayons pas $\theta_0 \in \Theta_{\delta^\pi}$ et donc que L_5 et L_1 (et L_2) aient des résultats différents.

2.2.2.5. Fonction de perte LINEX (L_6)

Rappelons que

$$\nu_{\delta^\pi(x)} \in \left(\frac{-1}{\alpha} \log (\mathbb{E}^{\pi(\theta|x)}[e^{-\alpha\theta}]) - \frac{3\varepsilon}{2} - \frac{2}{\alpha} \log \left(\frac{1 - e^{-\alpha\varepsilon}}{\alpha\varepsilon} \right); \right. \\ \left. \frac{-1}{\alpha} \log (\mathbb{E}^{\pi(\theta|x)}[e^{-\alpha\theta}]) - \frac{\varepsilon}{2} - \frac{2}{\alpha} \log \left(\frac{1 - e^{-\alpha\varepsilon}}{\alpha\varepsilon} \right) \right].$$

Puisque la distribution est asymétrique, il est impossible de déterminer la position de $\hat{\theta}_B = \frac{-1}{\alpha} \log (\mathbb{E}^{\pi(\theta|x)}[e^{-\alpha\theta}])$ par rapport à θ_0 .

2.3. DISTRIBUTION BIMODALE

2.3.1. Caractéristiques

Par distribution bimodale, nous entendons une distribution ayant deux modes (qui peuvent être des plateaux). La fonction de densité peut être ou n'être pas symétrique.

Si la distribution est symétrique, la moyenne (l'espérance) et la médiane ont la même valeur, mais elles ne correspondent pas aux modes que nous noterons θ_1 et θ_2 , avec $\theta_1 < \theta_2$. Si elle est asymétrique, nous ne pouvons tirer de conclusions. Un exemple de densité bimodale est un mélange de deux densités normales.

2.3.2. Résultats

2.3.2.1. Fonctions de perte 0-1 et 0-k (L_1 et L_2)

Rappelons que $\delta^\pi(x) = \arg \max_{a \in \mathbb{Z}} \Pr\{\theta \in \Theta_a | x\}$.

Nous présentons ici un avant-goût du théorème de la prochaine section. Même si nous n'avons que deux modes au lieu d'un nombre fini, les deux preuves sont très similaires.

Nous pouvons limiter à quatre le nombre d'intervalles possibles pour l'intervalle ayant la probabilité *a posteriori* maximale. Si nous notons les deux modes par $\theta_1 < \theta_2$ et que nous avons $\alpha_1 \leq \alpha_2$ tel que $\theta_1 \in \Theta_{\alpha_1}$ et $\theta_2 \in \Theta_{\alpha_2}$, alors Θ_{α_1} et Θ_{α_2} sont nos deux premiers candidats. Les deux autres sont voisins de ceux-ci.

Nous définissons les deux autres candidats comme suit :

$$\Theta_{\beta_1} = \begin{cases} \Theta_{\alpha_1-1} & \text{si } \pi(l_{\alpha_1}|x) > \pi(u_{\alpha_1}|x); \\ \Theta_{\alpha_1+1} & \text{si } \pi(l_{\alpha_1}|x) \leq \pi(u_{\alpha_1}|x); \end{cases}$$

$$\Theta_{\beta_2} = \begin{cases} \Theta_{\alpha_2-1} & \text{si } \pi(l_{\alpha_2}|x) \geq \pi(u_{\alpha_2}|x); \\ \Theta_{\alpha_2+1} & \text{si } \pi(l_{\alpha_2}|x) < \pi(u_{\alpha_2}|x). \end{cases}$$

La preuve se construit comme suit : en premier lieu, nous montrons qu'aucun intervalle à gauche du plus à gauche de Θ_{α_1} ou Θ_{β_1} ne peut être celui que nous recherchons. Ceci couvre en même temps la preuve du côté droit qui est similaire. Ensuite, il faut montrer que l'intervalle recherché ne peut se trouver parmi ceux qui ne sont pas sélectionnés entre Θ_{α_1} et Θ_{α_2} . Ceci requiert d'étudier tous les cas

selon les valeurs de β_1 et β_2 et selon la position du minimum de la densité (celui situé entre les deux modes).

2.3.2.2. Fonction de perte L_3

Rappelons que $\delta^\pi(x) = \arg \min_{a \in \mathbb{Z}} k_a (1 - \Pr\{\theta \in \Theta_a | x\})$.

Nous présentons le même résultat que pour la distribution unimodale asymétrique.

Notons par H_m l'hypothèse retenue par L_1 et L_2 . Alors $\delta^\pi(x) = m$ pour L_3 si $\nexists b \neq m$ tel que

$$k_b < k_m \frac{1 - \Pr\{\theta \in \Theta_m | x\}}{1 - \max_{a \in \{\alpha_1, \alpha_1 \pm 1, \alpha_2, \alpha_2 \pm 1\} \setminus \{m\}} \Pr\{\theta \in \Theta_a | x\}}.$$

Encore une fois, il est possible que ni θ_1 ni θ_2 ne fasse partie de Θ_d .

2.3.2.3. Fonction de perte L_4

Rappelons que $\delta^\pi(x) = \arg \max_{a \in \mathbb{Z}} k_a \Pr\{\theta \in \Theta_a | x\}$.

Ici aussi, nous nous contentons de présenter un résultat très similaire à celui de la distribution unimodale asymétrique.

Si nous notons H_m l'hypothèse retenue par L_1 et L_2 , nous aurons que $\delta^\pi(x) = m$ s'il n'existe pas d'entier b tel que

$$k_b > k_m \frac{\Pr\{\theta \in \Theta_m | x\}}{\max_{a \in \{\alpha_1, \alpha_1 \pm 1, \alpha_2, \alpha_2 \pm 1\} \setminus \{m\}} \Pr\{\theta \in \Theta_a | x\}}.$$

2.3.2.4. Fonction de perte quadratique (L_5)

Rappelons que $\delta^\pi(x) = \arg \min_{a \in \mathbb{Z}} (\nu_a - \mathbb{E}^{\pi(\theta|x)}[\theta])^2$.

Nous savons donc que $\mathbb{E}^{\pi(\theta|x)}[\theta] \in \Theta_{\delta^\pi}$, et il se peut très bien que le choix de L_5 ne s'accorde pas avec celui de L_1 et L_2 .

2.3.2.5. Fonction de perte LINEX (L_6)

Rappelons que

$$\nu_{\delta^\pi(x)} \in \left(\frac{-1}{\alpha} \log(\mathbb{E}^{\pi(\theta|x)}[e^{-\alpha\theta}]) - \frac{3\varepsilon}{2} - \frac{2}{\alpha} \log\left(\frac{1 - e^{-\alpha\varepsilon}}{\alpha\varepsilon}\right); \right. \\ \left. \frac{-1}{\alpha} \log(\mathbb{E}^{\pi(\theta|x)}[e^{-\alpha\theta}]) - \frac{\varepsilon}{2} - \frac{2}{\alpha} \log\left(\frac{1 - e^{-\alpha\varepsilon}}{\alpha\varepsilon}\right) \right].$$

Étant donné la grande variété de distributions possibles, nous ne pouvons tirer de conclusion intéressante.

2.4. DISTRIBUTION MULTIMODALE

2.4.1. Caractéristiques

Par distribution multimodale, nous entendons une distribution ayant une densité continue et un nombre fini de modes, mais au moins deux modes. Comme auparavant, les modes peuvent être des plateaux, auquel cas le mode est considéré comme étant le centre du plateau.

2.4.2. Résultat pour les fonctions de perte 0-1 et 0-k (L_1 et L_2)

Puisqu'il nous est impossible de trouver des résultats généraux pour les autres fonctions de perte, nous présentons ici une version simplifiée d'un théorème général sur le nombre d'intervalles qui peuvent être choisis par la décision de Bayes correspondant à L_1 et L_2 . Nous avons placé le théorème général en annexe.

Rappelons que pour L_1 et L_2 , $\delta^\pi(x) = \arg \max_{a \in \mathbb{Z}} \Pr\{\theta \in \Theta_a | x\}$.

Théorème 2.4.1. *Soit $\pi(\theta|x)$ la densité a posteriori d'une distribution à M modes, notés $\theta_1 < \theta_2 < \dots < \theta_M$.*

Définissons des intervalles Θ_i , $i \in \mathbb{Z}$ qui forment une partition de l'espace-paramètre Θ , avec les caractéristiques supplémentaires suivantes :

- $\Theta_i = (l_i, u_i]$, $u_i - l_i = \varepsilon > 0 \quad \forall i \in \mathbb{Z}$;
- $\Theta_{i+1} = (l_{i+1} = u_i, u_{i+1}]$, $\forall i \in \mathbb{Z}$.

Il existe M constantes $\alpha_1 \leq \alpha_2 \leq \dots \leq \alpha_M \in \mathbb{Z}$ tel que $\theta_i \in \Theta_{\alpha_i}$, $i = 1, 2, \dots, M$. Nous imposons maintenant la restriction que ε est suffisamment petit pour que $\alpha_{i+1} > \alpha_i + 1$, $i = 1, 2, \dots, M - 1$, et aussi pour qu'il n'y ait aucun minimum dans les intervalles Θ_{α_i} , $i = 1, 2, \dots, M$. Le théorème plus général abolit cette restriction qui simplifie beaucoup la preuve.

Il est possible de définir un β_i pour chaque α_i de telle façon que

$$\arg \max_{a \in \mathbb{Z}} \Pr\{\theta \in \Theta_a | x\} \in \{\alpha_1, \alpha_2, \dots, \alpha_M, \beta_1, \beta_2, \dots, \beta_M\}.$$

Les β_i sont définis ainsi :

$$\beta_i = \begin{cases} \alpha_i + 1 & \text{si } \pi(l_{\alpha_i} | x) < \pi(u_{\alpha_i} | x); \\ \alpha_i - 1 & \text{si } \pi(l_{\alpha_i} | x) > \pi(u_{\alpha_i} | x) \text{ et } \beta_{i-1} \neq \alpha_i - 1; \\ \text{pas défini sinon.} \end{cases}$$

Notons que β_1 et β_M sont définis différemment dans le théorème général.

Démonstration (du théorème 2.4.1) :

Nous commençons par montrer que

$$\arg \max_{a \in \mathbb{Z}} \Pr\{\theta \in \Theta_a | x\} \not\prec \min\{\alpha_1, \beta_1\}.$$

Puisqu'il n'y a pas de mode à gauche de θ_1 , nous savons que

$$\pi(s|x) \leq \pi(t|x) \leq \pi(\theta_1|x), \forall s \leq t \leq \theta_1.$$

Alors, dans le cas où $\beta_1 = \alpha_1 - 1$, nous avons que

$$\begin{aligned} \Pr\{\theta \in \Theta_a | x\} &\leq \varepsilon \pi(u_a | x) \\ &\leq \varepsilon \pi(l_{\beta_1} | x) \\ &\leq \Pr\{\theta \in \Theta_{\beta_1} | x\}, \forall a < \beta_1. \end{aligned}$$

Dans le cas où $\beta_1 \neq \alpha_1 - 1$, c'est-à-dire que $\pi(l_{\alpha_1}|x) \leq \pi(u_{\alpha_1}|x)$,

$$\begin{aligned} \Pr\{\theta \in \Theta_a|x\} &\leq \varepsilon\pi(u_a|x) \\ &\leq \varepsilon\pi(l_{\alpha_1}|x) \\ &\leq \Pr\{\theta \in \Theta_{\alpha_1}|x\}, \forall a < \alpha_1. \end{aligned}$$

Notons que $\varepsilon\pi(l_{\alpha_1}|x) \leq \Pr\{\theta \in \Theta_{\alpha_1}|x\}$ est valide parce que nous avons imposé la restriction qu'il n'y ait pas de minimum dans les intervalles Θ_{α_i} . Sinon, il se peut que $\pi(l_{\alpha_1}|x) \leq \pi(u_{\alpha_1}|x)$ et $\varepsilon\pi(l_{\alpha_1}|x) > \Pr\{\theta \in \Theta_{\alpha_1}|x\}$.

Nous prouvons de façon similaire que

$$\arg \max_{a \in \mathbb{Z}} \Pr\{\theta \in \Theta_a|x\} \neq \max\{\alpha_M, \beta_M\}.$$

Il ne nous reste qu'à montrer que

$$\arg \max_{a \in \mathbb{Z}} \Pr\{\theta \in \Theta_a|x\} \notin \{M_i + 1, M_i + 2, \dots, m_{i+1} - 2, m_{i+1} - 1\}$$

où $M_i = \max\{\alpha_i, \beta_i\}$ et $m_{i+1} = \min\{\alpha_{i+1}, \beta_{i+1}\}$, $i = 1, 2, \dots, M - 1$.

Nous devons séparer les cas selon que $M_i = \alpha_i$ ou $M_i = \beta_i$, ainsi que $m_{i+1} = \alpha_{i+1}$ ou $m_{i+1} = \beta_{i+1}$. Nous n'étudions que les cas où $\{M_i + 1, M_i + 2, \dots, m_{i+1} - 2, m_{i+1} - 1\} \neq \emptyset$, les autres étant triviaux.

- 1^{er} cas : $M_i = \alpha_i$ et $m_{i+1} = \alpha_{i+1}$

Rappelons que $M_i = \alpha_i \Leftrightarrow \pi(l_{\alpha_i}|x) \geq \pi(u_{\alpha_i}|x)$ et $m_{i+1} = \alpha_{i+1} \Leftrightarrow \pi(l_{\alpha_{i+1}}|x) \leq \pi(u_{\alpha_{i+1}}|x)$ par définition. Nous avons donc que $\Pr\{\theta \in \Theta_{\alpha_i}|x\} \geq \varepsilon\pi(u_{\alpha_i}|x)$ et $\Pr\{\theta \in \Theta_{\alpha_{i+1}}|x\} \geq \varepsilon\pi(l_{\alpha_{i+1}}|x)$.

Il existe un point $\theta^* \in (u_{\alpha_i}, l_{\alpha_{i+1}})$ (le minimum) tel que

$$\pi(u_{\alpha_i}|x) \geq \pi(s|x) \geq \pi(t|x) \geq \pi(\theta^*|x), u_{\alpha_i} \leq s \leq t \leq \theta^*$$

et

$$\pi(\theta^*|x) \leq \pi(s|x) \leq \pi(t|x) \leq \pi(l_{\alpha_{i+1}}|x), \theta^* \leq s \leq t \leq l_{\alpha_{i+1}}.$$

Nous avons donc que

$$\pi(\theta|x) \leq \max\{\pi(u_{\alpha_i}|x); \pi(l_{\alpha_{i+1}}|x)\} \quad \forall \theta \in (u_{\alpha_i}, l_{\alpha_{i+1}}).$$

Ceci nous permet d'affirmer que $\forall a$ tel que $\alpha_i < a < \alpha_{i+1}$,

$$\begin{aligned} \Pr\{\theta \in \Theta_a|x\} &\leq \varepsilon \max\{\pi(u_{\alpha_i}|x); \pi(l_{\alpha_{i+1}}|x)\} \\ &\leq \max\{\Pr\{\theta \in \Theta_{\alpha_i}|x\}; \Pr\{\theta \in \Theta_{\alpha_{i+1}}|x\}\}. \end{aligned}$$

– 2^e cas : $M_i = \alpha_i$ et $m_{i+1} = \beta_{i+1}$

Comme précédemment, $\Pr\{\theta \in \Theta_{\alpha_i}|x\} \geq \varepsilon\pi(u_{\alpha_i}|x)$. Aussi, $\exists \theta^* \in (u_{\alpha_i}, l_{\alpha_{i+1}})$ avec les mêmes caractéristiques. Nous devons étudier séparément les deux possibilités : $\theta^* \in \Theta_{\beta_{i+1}}$ ou $\theta^* \notin \Theta_{\beta_{i+1}}$.

– $\theta^* \notin \Theta_{\beta_{i+1}}$

La preuve revient à celle du premier cas, à la seule différence que nous remplaçons les α_{i+1} par β_{i+1} .

– $\theta^* \in \Theta_{\beta_{i+1}}$

Nous avons alors que $\pi(\theta|x) \leq \pi(u_{\alpha_i}|x) \quad \forall \theta \in (u_{\alpha_i}, l_{\beta_{i+1}})$.

Ceci nous permet d'affirmer que $\forall a$ tel que $\alpha_i < a < \beta_{i+1}$,

$$\begin{aligned} \Pr\{\theta \in \Theta_a|x\} &\leq \varepsilon\pi(l_a|x) \\ &\leq \varepsilon\pi(u_{\alpha_i}|x) \\ &\leq \Pr\{\theta \in \Theta_{\alpha_i}|x\}. \end{aligned}$$

– 3^e cas : $M_i = \beta_i$ et $m_{i+1} = \alpha_{i+1}$

Ce cas est symétrique au 2^e cas, donc il n'est pas nécessaire de faire une preuve qui serait redondante.

– 4^e cas : $M_i = \beta_i$ et $m_{i+1} = \beta_{i+1}$

Nous séparons en 3 cas selon la position de θ^* , qui est défini comme précédemment.

– $\theta^* \notin \Theta_{\beta_i}$ et $\theta^* \notin \Theta_{\beta_{i+1}}$

Nous retournons au 1^{er} cas, en remplaçant α_i par β_i et α_{i+1} par β_{i+1} .

– $\theta^* \in \Theta_{\beta_{i+1}}$

Nous retournons à la première possibilité du 2^e cas, en remplaçant α_i par β_i .

– $\theta^* \in \Theta_{\beta_i}$

Ce cas est symétrique au cas précédent.

2.5. CONCLUSION

Nous avons appliqué les résultats obtenus au chapitre 1 pour les fonctions de perte et les décisions de Bayes y correspondant à quelques types de distributions pour obtenir quelques résultats plus spécifiques.

Pour L_1 et L_2 , nous avons obtenu des résultats pour trouver l'intervalle ayant la probabilité *a posteriori* maximale à partir des modes de la distribution. Pour L_3 et L_4 , étant donné que les résultats dépendent beaucoup des constantes, nous avons obtenu quelques résultats sur les relations entre les constantes.

Pour L_5 , la décision de Bayes correspond toujours à l'intervalle contenant l'espérance *a posteriori*, ce qui simplifie les résultats. Il correspond à celui contenant le mode pour les distributions unimodales symétriques car dans ce cas le mode et l'espérance (lorsqu'elle existe) coïncident. Pour L_6 , notre résultat qui nous donne l'intervalle contenant le centre de l'intervalle retenu tient toujours et nous obtenons un résultat supplémentaire pour les distributions unimodales symétriques.

Finalement, nous avons présenté un théorème pour les fonctions de perte L_1 et L_2 qui nous donne le nombre et exactement quels intervalles peuvent être celui ayant la probabilité *a posteriori* maximale pour une distribution multimodale. Ce théorème requiert certaines restrictions sur la taille de la partition ε , qui sont relâchées dans le théorème qui se retrouve à l'annexe.

Chapitre 3

COMPARAISON AVEC L'ESTIMATION PONCTUELLE, SIMULATIONS ET EXEMPLE

Nous abordons maintenant le dernier chapitre, où nous commençons par comparer les résultats obtenus par les tests d'hypothèses multiples et les estimateurs ponctuels correspondants. Nous examinons aussi ce qui se passe lorsque nous faisons diminuer la largeur des intervalles vers 0.

Nous présentons ensuite les résultats des simulations faites pour illustrer les résultats de ce chapitre. Nous finirons en présentant les résultats que nous avons obtenus en travaillant sur des données réelles provenant des montants de réclamation d'assurance-incendie d'une compagnie d'assurance danoise en 1981.

3.1. COMPARAISON AVEC L'ESTIMATION PONCTUELLE

Nous gardons le même ordre qu'aux chapitres 1 et 2 pour la présentation des fonctions de perte et de leurs résultats.

3.1.1. Fonctions de perte à valeurs constantes

3.1.1.1. Fonctions de perte 0-1 et 0-k (L_1 et L_2)

L'estimateur ponctuel correspondant à la fonction de perte 0-1 est l'estimateur MAP (maximum *a posteriori*), que nous noterons $\hat{\theta}_{MAP}$ et qui est défini par

$$\hat{\theta}_{MAP} = \arg \max_{\theta \in \mathbb{R}} \pi(\theta|x).$$

En fait, dans Robert(2001), l'estimateur MAP est défini comme la limite des estimateurs de Bayes de la fonction de perte $L_\varepsilon(d, \theta) = I_{|\theta-d|>\varepsilon}$ lorsque $\varepsilon \rightarrow 0$, la fonction $I_{|\theta-d|>\varepsilon}$ prenant la valeur 1 si $|\theta - d| > \varepsilon$ et 0 sinon.

Malheureusement, sauf dans les cas où la distribution est symétrique et qu'il existe un mode global unique, il nous est impossible de définir une suite de décisions qui converge vers la décision telle que le mode global (unique) est dans l'intervalle de l'hypothèse retenue, lorsque $\varepsilon \rightarrow 0$, ce qui implique qu'il n'y a pas toujours accord entre les tests d'hypothèses multiples et l'estimation ponctuelle.

La raison de ceci est que dans le cas de l'estimateur MAP, les estimateurs de Bayes peuvent prendre toutes les valeurs dans $\mathbb{R} \forall \varepsilon$, alors que les intervalles de notre partition sont fixes pour chaque ε . Pour certains ε nous choisissons l'hypothèse telle que le mode global (unique) est dans l'intervalle correspondant, mais ce ne sera pas nécessairement le cas pour tous les ε lorsque $\varepsilon \rightarrow 0$ (à moins que la distribution soit symétrique). Ces résultats ont été obtenus aux sous-sections 2.1.2.1, 2.2.2.1 et 2.3.2.1.

Malgré tout, les résultats pour les fonctions de perte L_1 et L_2 que nous avons obtenus aux sous-sections mentionnées précédemment tiennent toujours, puisque $\varepsilon > 0$. De plus, pour les distributions (nous assumons encore qu'elles sont continues) ayant deux modes ou plus, il est parfois possible de simplifier les résultats et de limiter encore plus le nombre d'intervalles pouvant être celui ayant la probabilité *a posteriori* maximale.

En effet, dû à la continuité de la fonction de densité, si la densité à l'un des modes est supérieure à celle des autres, les intervalles α_i et β_i correspondants à ce mode ne pourront pas être choisis pour ε suffisamment petit. Nous pouvons utiliser ce résultat à répétition pour trouver ε tel qu'il ne nous restera que le mode global (ou les modes globaux, s'il y en a plusieurs), ainsi que les intervalles

adjacents à ceux-ci, comme possibilités.

Proposition 3.1.1. *Soit $\pi(\theta|x)$ la fonction de densité continue d'une distribution ayant M modes, que nous noterons $\theta_1, \theta_2, \dots, \theta_M$, tel que $\theta_1 < \theta_2 < \dots < \theta_M$, $M \geq 2$.*

Classons les modes sous la notation $\theta_{(1)}, \theta_{(2)}, \dots, \theta_{(M)}$ tel que

$$\pi(\theta_{(1)}|x) \leq \pi(\theta_{(2)}|x) \leq \dots \leq \pi(\theta_{(M)}|x).$$

Supposons que

$$\begin{aligned} \pi(\theta_{(1)}|x) &= \dots = \pi(\theta_{(I)}|x) \\ &< \pi(\theta_{(I+1)}|x) \leq \dots \leq \pi(\theta_{(M)}|x), \quad 1 \leq I < M. \end{aligned}$$

Alors nous avons que

$$\begin{aligned} &\max_{a \in \{\alpha_{(1)}, \alpha_{(2)}, \dots, \alpha_{(I)}, \beta_{(1)}, \beta_{(2)}, \dots, \beta_{(I)}\}} \Pr\{\theta \in \Theta_a|x\} \\ &\leq \max_{a \in \{\alpha_{(I+1)}, \dots, \alpha_{(M)}\}} \Pr\{\theta \in \Theta_a|x\} \end{aligned}$$

si nous choisissons $\varepsilon < \max\{\varepsilon_{(I+1)}, \varepsilon_{(I+2)}, \dots, \varepsilon_{(M)}\}$ où

$$\varepsilon_{(i)} = \inf\{t \in R^+ : \min\{\pi(\theta_{(i)} - t|x); \pi(\theta_{(i)} + t|x)\} < \pi(\theta_{(1)}|x)\},$$

$$i = I + 1, I + 2, \dots, M.$$

Démonstration (proposition 3.1.1) :

Puisque $\varepsilon < \max\{\varepsilon_{(I+1)}, \varepsilon_{(I+2)}, \dots, \varepsilon_{(M)}\}$, $\exists j \in \{I + 1, I + 2, \dots, M\}$ tel que $\varepsilon < \varepsilon_{(j)}$. D'après la définition des $\varepsilon_{(i)}$, ceci implique que

$$\pi(t|x) \geq \pi(\theta_{(1)}|x) \quad \forall t \in (\theta_{(j)} - \varepsilon, \theta_{(j)} + \varepsilon).$$

Or, nous savons que $\theta_{(j)} \in \Theta_{\alpha_{(j)}}$, ce qui implique que $\Theta_{\alpha_{(j)}} \subset (\theta_{(j)} - \varepsilon, \theta_{(j)} + \varepsilon]$. Nous avons donc le résultat

$$\begin{aligned} \Pr\{\theta \in \Theta_{\alpha_{(j)}}|x\} &\geq \varepsilon\pi(\theta_{(1)}|x) \\ &\geq \max_{a \in \{\alpha_{(1)}, \alpha_{(2)}, \dots, \alpha_{(I)}, \beta_{(1)}, \beta_{(2)}, \dots, \beta_{(I)}\}} \Pr\{\theta \in \Theta_a|x\} \end{aligned}$$

car $\pi(\theta_{(1)}|x) = \pi(\theta_{(2)}|x) = \dots = \pi(\theta_{(I)}|x)$.

En fait, il est possible qu'il existe un entier $k \in \{1, 2, \dots, I\}$ tel que nous avons $\max_{a \in \{\alpha_{(k)}, \beta_{(k)}\}} \Pr\{\theta \in \Theta_a|x\} > \varepsilon\pi(\theta_{(k)}|x)$ mais ceci implique qu'il y a un minimum dans $\Theta_{\alpha_{(k)}}$ ou $\Theta_{\beta_{(k)}}$. Mais, alors, si le minimum en question est celui à gauche de $\theta_{(k)}$, nous avons que

$$\max_{a \in \{\alpha_{(k)-1}, \beta_{(k)-1}\}} \Pr\{\theta \in \Theta_a|x\} \geq \max_{a \in \{\alpha_{(k)}, \beta_{(k)}\}} \Pr\{\theta \in \Theta_a|x\}.$$

Si ceci n'est pas vérifié, alors nous pouvons appliquer le lemme de l'annexe à ce résultat. Pour une preuve plus détaillée, voir le cas où le minimum (noté θ^*) est dans Θ_{α_i} ou un intervalle adjacent dans la troisième partie (celle qui concerne les intervalles entre deux modes) de la preuve du théorème en annexe. Notons qu'il faut ajouter une condition : si le minimum est à gauche, nous aurons $\pi(\min\{l_{\alpha_i}, l_{\beta_i}\}|x) > \pi(\theta_i|x)$; si le minimum est à droite, c'est $\pi(\max\{u_{\alpha_i}, u_{\beta_i}\}|x) > \pi(\theta_i|x)$.

Le résultat de cette proposition est que nous ne retenons pas les intervalles $\{\Theta_{\alpha_{(1)}}, \Theta_{\alpha_{(2)}}, \dots, \Theta_{\alpha_{(I)}}, \Theta_{\beta_{(1)}}, \Theta_{\beta_{(2)}}, \dots, \Theta_{\beta_{(I)}}\}$ comme candidats pour l'intervalle ayant la probabilité *a posteriori* maximale.

Proposition 3.1.2. Soit $\pi(\theta|x)$ la fonction de densité continue d'une distribution ayant M modes, que nous notons $\theta_1, \theta_2, \dots, \theta_M$, tel que $\theta_1 < \theta_2 < \dots < \theta_M$. $M \geq 2$.

Classons les modes sous la notation $\theta_{(1)}, \theta_{(2)}, \dots, \theta_{(M)}$ tel que

$$\pi(\theta_{(1)}|x) \leq \pi(\theta_{(2)}|x) \leq \dots \leq \pi(\theta_{(M)}|x).$$

Supposons que

$$\begin{aligned} \pi(\theta_{(1)}|x) &\leq \pi(\theta_{(2)}|x) \leq \dots \leq \pi(\theta_{(I)}|x) \\ &< \pi(\theta_{(I+1)}|x) = \pi(\theta_{(I+2)}|x) = \dots = \pi(\theta_{(M)}|x), 1 < I + 1 \leq M. \end{aligned}$$

Nous avons alors que

$$\begin{aligned} &\max_{a \in \{\alpha_{(1)}, \alpha_{(2)}, \dots, \alpha_{(I)}, \beta_{(1)}, \beta_{(2)}, \dots, \beta_{(I)}\}} \Pr\{\theta \in \Theta_a | x\} \\ &\leq \max_{\{a \in \alpha_{(I+1)}, \alpha_{(I+2)}, \dots, \alpha_{(M)}\}} \Pr\{\theta \in \Theta_a | x\} \end{aligned}$$

si nous choisissons $\varepsilon < \max\{\varepsilon_{(I+1)}, \varepsilon_{(I+2)}, \dots, \varepsilon_{(M)}\}$ où

$$\varepsilon_{(i)} = \inf\{t \in R^+ : \min\{\pi(\theta_{(i)} - t|x); \pi(\theta_{(i)} + t|x)\} < \pi(\theta_{(I)}|x)\},$$

$$i = I + 1, I + 2, \dots, M.$$

Démonstration (proposition 3.1.2) :

La preuve revient pratiquement à la même chose que pour la proposition précédente, puisque $\pi(\theta_{(1)}|x) \leq \pi(\theta_{(2)}|x) \leq \dots \leq \pi(\theta_{(M)}|x)$.

Le résultat de cette proposition est que seuls les intervalles dont les modes sont les plus élevés, soit $\{\Theta_{\alpha_{(I+1)}}, \Theta_{\alpha_{(I+2)}}, \dots, \Theta_{\alpha_{(M)}}, \Theta_{\beta_{(I+1)}}, \Theta_{\beta_{(I+2)}}, \dots, \Theta_{\beta_{(M)}}\}$, peuvent être l'intervalle ayant la probabilité *a posteriori* maximale.

Nous pouvons retenir de cette sous-section que les tests d'hypothèses multiples et l'estimation ponctuelle ne s'accordent pas nécessairement, à cause du fait que les centres des intervalles sont fixés d'avance pour les tests d'hypothèses multiples, alors qu'ils ne le sont pas dans la méthode présentée par Robert(2001) pour obtenir l'estimateur ponctuel correspondant à L_1 et L_2 .

3.1.1.2. Fonction de perte L_3

Il n'existe pas comme tel d'estimateur ponctuel correspondant à L_3 et donc une comparaison exacte est impossible.

Par contre, nous savons que, pour L_3 ,

$$\delta^\pi(x) = \arg \min_{a \in \mathbb{Z}} k_a(1 - \Pr\{\theta \in \Theta_a|x\}).$$

Mais, pour une distribution dont la fonction de densité est continue et bornée, nous avons que $\forall a, \Pr\{\theta \in \Theta_a|x\} \rightarrow 0$ lorsque $\varepsilon \rightarrow 0$. Donc nous avons que

$$\delta^\pi(x) \xrightarrow{\varepsilon \rightarrow 0} \arg \min_{a \in \mathbb{Z}} k_a.$$

Dans un cas d'estimation ponctuelle, nous pourrions définir

$$L_3^*(\theta, d) = \begin{cases} 0 & \text{si } d = \theta; \\ k(d) & \text{si } d \neq \theta. \end{cases}$$

Nous choisirions alors $\delta^\pi(x) = \arg \min_{d \in \mathbb{R}} k(d)$. C'est un résultat peu intéressant puisque la distribution *a posteriori* n'a aucune influence sur la décision, mais on voit que pour ces deux définitions, l'estimation ponctuelle et les tests d'hypothèses multiples s'accordent.

3.1.1.3. Fonction de perte L_4

Encore une fois, il n'existe pas d'estimateur ponctuel correspondant exactement à notre fonction de perte, mais nous pouvons définir une fonction de perte très similaire comme suit :

$$L_4^*(\theta, d) = \begin{cases} 0 & \text{si } d = \theta; \\ k(\theta) & \text{si } d \neq \theta. \end{cases}$$

L'estimateur ponctuel de L_4^* sera

$$\delta^\pi(x) = \arg \max_{d \in \mathbb{R}} k(d)\pi(d|x).$$

Si nous ajoutons la condition que la fonction $k(d)$ soit continue ainsi que $\lim_{d \rightarrow \pm\infty} k(d)\pi(d|x) = 0$, nous pouvons retourner à la section 3.1.1.1 en utilisant $\pi^*(\theta|x) = k(\theta)\pi(\theta|x)$ au lieu de $\pi(\theta|x)$. Ce résultat est similaire au résultat présenté par Robert(2001) pour une fonction de perte quadratique pondérée, où il fait la remarque qu'une fonction de perte quadratique pondérée par $w(\theta)$ donne le même résultat qu'une fonction de perte quadratique (non pondérée) avec une

densité *a priori* repondérée $\pi_w(\theta) \propto \pi(\theta)w(\theta)$.

Nous pouvons donc conclure que les tests d'hypothèses multiples et l'estimation ponctuelle avec L_4^* ne s'accordent pas nécessairement, comme nous l'avons vu à la sous-section 3.1.1.1.

3.1.2. Fonctions de perte à valeurs continues

3.1.2.1. Fonction de perte quadratique (L_5)

La fonction de perte correspondante pour l'estimation ponctuelle est la même, mais elle n'a pas à être intégrée sur un intervalle. Nous trouvons comme estimateur ponctuel

$$\delta^\pi(x) = \mathbb{E}^{\pi(\theta|x)}[\theta].$$

Or, nous avons trouvé que l'hypothèse retenue par un test d'hypothèses multiples est celle tel que

$$\begin{aligned} \delta^\pi(x) &= \arg \min_{a \in \mathbb{Z}} (\nu_a - \mathbb{E}^{\pi(\theta|x)}[\theta])^2 \\ &= \arg \min_{a \in \mathbb{Z}} |\nu_a - \mathbb{E}^{\pi(\theta|x)}[\theta]| \end{aligned}$$

où ν_a est le centre de l'intervalle Θ_a .

Cette hypothèse est donc celle dont le centre est le plus près de $\mathbb{E}^{\pi(\theta|x)}[\theta]$. De ce fait, nous savons que l'estimateur ponctuel sera toujours dans l'intervalle choisi.

3.1.2.2. Fonction de perte LINEX (L_6)

Comme pour L_5 , la fonction de perte correspondante pour l'estimation ponctuelle est la même, mais sans intégrale. Nous obtenons comme estimateur ponctuel

$$\delta^\pi(x) = \hat{\theta}_B = \frac{-1}{\alpha} \log (\mathbb{E}^{\pi(\theta|x)}[e^{-\alpha\theta}])$$

comme nous l'avons dans le lemme de la section 1.3.2 sur la fonction de perte LINEX. Notons que certaines distributions n'ont pas d'estimateur LINEX, comme

par exemple une distribution *a posteriori* lognormale.

Nous savons aussi que le test d'hypothèses multiples choisira l'hypothèse telle que

$$\nu_{\delta^\pi(x)} \in \left(\frac{-1}{\alpha} \log(\mathbb{E}^{\pi(\theta|x)}[e^{-\alpha\theta}]) - \frac{3\varepsilon}{2} - \frac{2}{\alpha} \log\left(\frac{1 - e^{-\alpha\varepsilon}}{\alpha\varepsilon}\right); \right. \\ \left. \frac{-1}{\alpha} \log(\mathbb{E}^{\pi(\theta|x)}[e^{-\alpha\theta}]) - \frac{\varepsilon}{2} - \frac{2}{\alpha} \log\left(\frac{1 - e^{-\alpha\varepsilon}}{\alpha\varepsilon}\right) \right].$$

Proposition 3.1.3. *Pour $\alpha\varepsilon \rightarrow 0$, nous avons que le test d'hypothèses multiples choisira l'hypothèse $H_{\delta^\pi(x)}$ tel que*

$$\nu_{\delta^\pi(x)} \in \left(\hat{\theta}_B - \frac{\varepsilon}{2} + O(\varepsilon^2); \hat{\theta}_B + \frac{\varepsilon}{2} + O(\varepsilon^2) \right].$$

Démonstration (proposition 3.1.3) :

Nous savons que

$$e^{-\alpha\varepsilon} = 1 - \alpha\varepsilon + \frac{(\alpha\varepsilon)^2}{2} - \frac{(\alpha\varepsilon)^3}{6} + \dots$$

et donc que

$$\frac{1 - e^{-\alpha\varepsilon}}{\alpha\varepsilon} = 1 - \frac{\alpha\varepsilon}{2} + \frac{(\alpha\varepsilon)^2}{6} - \dots \\ = 1 - \frac{\alpha\varepsilon}{2} + O((\alpha\varepsilon)^2).$$

De plus, $\log(1 + x) = x + O(x^2)$ (pour x très petit) et donc

$$\log\left(\frac{1 - e^{-\alpha\varepsilon}}{\alpha\varepsilon}\right) = -\frac{\alpha\varepsilon}{2} + O((\alpha\varepsilon)^2).$$

Nous arrivons finalement à

$$-\frac{\varepsilon}{2} - \frac{2}{\alpha} \log\left(\frac{1 - e^{-\alpha\varepsilon}}{\alpha\varepsilon}\right) = \frac{\varepsilon}{2} + O(\varepsilon^2)$$

ce qui nous donne notre résultat.

Nous voyons donc que plus $\alpha\varepsilon$ est petit, plus nous approchons du résultat

$$\nu_{\delta^\pi(x)} \in \left(\hat{\theta}_B - \frac{\varepsilon}{2}; \hat{\theta}_B + \frac{\varepsilon}{2} \right].$$

Ceci veut dire que lorsque $\alpha\varepsilon \rightarrow 0$ (et puisque $\alpha > 0$, nous utiliserons $\varepsilon \rightarrow 0$), nous choisissons presque toujours l'hypothèse dont l'intervalle contient l'estimateur ponctuel. Nous avons donc que l'estimateur ponctuel s'accorde avec la limite (lorsque $\varepsilon \rightarrow 0$) du test d'hypothèses multiples.

Par contre, ce n'est pas nécessairement le cas lorsque $\alpha\varepsilon$ ne tend pas vers 0. Par exemple, si nous avons $\alpha \rightarrow \infty$ (pour que $\alpha\varepsilon \rightarrow \infty$, puisque $\varepsilon < \infty$), alors nous obtenons que pour le test d'hypothèses multiples, l'intervalle devient

$$\nu_{\delta^\pi(x)} \in \left(\hat{\theta}_B - \frac{3\varepsilon}{2}; \hat{\theta}_B - \frac{\varepsilon}{2} \right],$$

ce qui implique que l'intervalle retenu ne contiendra jamais l'estimateur ponctuel.

Si $\alpha \rightarrow -\infty$, nous avons alors que le résultat tendra vers

$$\nu_{\delta^\pi(x)} \in \left(\hat{\theta}_B + \frac{\varepsilon}{2}; \hat{\theta}_B + \frac{3\varepsilon}{2} \right].$$

Pour avoir des résultats concrets pour des cas moins extrêmes (donc $0 < |\alpha\varepsilon| < \infty$), nous devons déterminer une distribution pour les centres des intervalles. En fait, nous n'avons à déterminer qu'un des centres puisque les autres sont tous situés à une distance de lui qui est un multiple de ε . Nous supposons que la distribution du centre placé dans l'intervalle $(0, \varepsilon)$ sera uniforme sur cet intervalle.

Proposition 3.1.4. *Si la distribution du centre d'un intervalle est uniforme sur le support \mathbb{R} , nous avons que*

$$\Pr\{\hat{\theta}_B \in \Theta_{\delta^\pi(x)}|x\} = \begin{cases} \frac{-2}{\alpha\varepsilon} \log\left(\frac{1 - e^{-\alpha\varepsilon}}{\alpha\varepsilon}\right) & \text{si } \alpha > 0; \\ 2 + \frac{2}{\alpha\varepsilon} \log\left(\frac{1 - e^{-\alpha\varepsilon}}{\alpha\varepsilon}\right) & \text{si } \alpha < 0. \end{cases}$$

Avant de passer à la preuve de cette proposition, notons que le côté droit donne toujours un résultat entre 0 et 1, comme il se doit pour une probabilité. Notons aussi que nous pourrions utiliser seulement la partie faite pour $\alpha > 0$ puisque les valeurs dépendent seulement de la valeur de $|\alpha\varepsilon|$ et pas de son signe.

Démonstration (proposition 3.1.4) :

Puisque les centres d'intervalle sont distribués de façon uniforme, la densité de $\nu_{\delta^\pi(x)}$ est

$$g_{\delta^\pi(x)}(v|\alpha\varepsilon) = \begin{cases} \frac{1}{\varepsilon} \text{ si } v \in \left(\hat{\theta}_B - \frac{3\varepsilon}{2} - \frac{2}{\alpha} \log \left(\frac{1 - e^{-\alpha\varepsilon}}{\alpha\varepsilon} \right); \right. \\ \left. \hat{\theta}_B - \frac{\varepsilon}{2} - \frac{2}{\alpha} \log \left(\frac{1 - e^{-\alpha\varepsilon}}{\alpha\varepsilon} \right) \right]; \\ 0 \text{ sinon.} \end{cases}$$

Nous savons de plus que

$$\hat{\theta}_B \in \Theta_{\delta^\pi(x)} \Leftrightarrow \nu_{\delta^\pi(x)} \in \left(\hat{\theta}_B - \frac{\varepsilon}{2}; \hat{\theta}_B + \frac{\varepsilon}{2} \right].$$

Puisque pour $\alpha > 0$, $-\frac{\varepsilon}{2} - \frac{2}{\alpha} \log \left(\frac{1 - e^{-\alpha\varepsilon}}{\alpha\varepsilon} \right) < \frac{\varepsilon}{2}$, nous trouvons

$$\begin{aligned} \Pr\{\hat{\theta}_B \in \Theta_{\delta^\pi(x)}|x\} &= \Pr\left\{\nu_{\delta^\pi(x)} \in \left(\hat{\theta}_B - \frac{\varepsilon}{2}; \hat{\theta}_B + \frac{\varepsilon}{2} \right) | x\right\} \\ &= \int_{\hat{\theta}_B - \frac{\varepsilon}{2}}^{\hat{\theta}_B + \frac{\varepsilon}{2}} g_{\nu_{\delta^\pi(x)}}(t|\alpha\varepsilon) dt \\ &= \int_{\hat{\theta}_B}^{\hat{\theta}_B - \frac{\varepsilon}{2} - \frac{2}{\alpha} \log \left(\frac{1 - e^{-\alpha\varepsilon}}{\alpha\varepsilon} \right)} \frac{1}{\varepsilon} dt \\ &= \frac{-2}{\alpha\varepsilon} \log \left(\frac{1 - e^{-\alpha\varepsilon}}{\alpha\varepsilon} \right). \end{aligned}$$

Pour $\alpha < 0$, $-\frac{\varepsilon}{2} - \frac{2}{\alpha} \log \left(\frac{1 - e^{-\alpha\varepsilon}}{\alpha\varepsilon} \right) > \frac{\varepsilon}{2}$ et nous trouvons

$$\begin{aligned} \Pr\{\theta \in \Theta_{\delta^\pi(x)}|x\} &= \int_{\hat{\theta}_B - \frac{3\varepsilon}{2} - \frac{2}{\alpha} \log \left(\frac{1 - e^{-\alpha\varepsilon}}{\alpha\varepsilon} \right)}^{\hat{\theta}_B + \frac{\varepsilon}{2}} \frac{1}{\varepsilon} dt \\ &= 2 + \frac{2}{\alpha\varepsilon} \log \left(\frac{1 - e^{-\alpha\varepsilon}}{\alpha\varepsilon} \right). \end{aligned}$$

Il est important de noter que ce résultat ne donne qu'une probabilité et ne permet aucunement de déterminer si $\hat{\theta}_B \in \Theta_{\delta^\pi(x)}$. Notons toutefois que pour $\alpha > 0$, soit $\hat{\theta}_B \in \Theta_{\delta^\pi(x)}$, soit $\hat{\theta}_B \in \Theta_{\delta^\pi(x)+1}$. Similairement, pour $\alpha < 0$, soit $\hat{\theta}_B \in \Theta_{\delta^\pi(x)}$, soit $\hat{\theta}_B \in \Theta_{\delta^\pi(x)-1}$. Finalement, la probabilité est la même des deux

côtés pour une même valeur de $|\alpha\varepsilon|$, puisqu'on ne fait qu'inverser l'orientation de la LINEX et que la distribution du paramètre n'est pas impliquée.

3.2. SIMULATIONS

Nous présentons dans cette section les résultats de nos simulations visant à illustrer et possiblement valider les résultats théoriques de nos deux premiers chapitres et de la sous-section 3.1.2.2 sur la fonction de perte LINEX (L_6).

Nous commençons par montrer nos résultats de simulations sur l'accord entre l'estimation ponctuelle et les tests d'hypothèses multiples, pour les fonctions de perte L_1 (donc L_2 aussi), quadratique (L_5) et LINEX (L_6) avec $\alpha = \frac{1}{2}, 1, 2$. Ensuite, nous illustrons le résultat de la sous-section 3.1.2.2 sur la probabilité que l'intervalle retenu par le test d'hypothèses multiples contienne l'estimateur ponctuel pour la LINEX, probabilité qui dépend de la valeur de $\alpha\varepsilon$.

Finalement, nous illustrons les résultats des pertes espérées pour toutes les fonctions de perte et tous les intervalles pour chacun des trois types de distribution, avec un exemple pour chacun des types.

3.2.1. Accord entre estimation ponctuelle et test d'hypothèses multiples

Les résultats des 10 000 simulations se retrouvent au tableau 3.2.1. Nous avons choisi une normale comme distribution *a posteriori* pour représenter le type unimodal symétrique, une gamma pour le type unimodal asymétrique et un mélange 80%-20% de deux normales pour le type bimodal. À chaque fois, nous avons généré un échantillon de 500 observations. De plus, afin d'éviter d'obtenir approximativement les mêmes intervalles (et les mêmes distributions) à chaque simulation, nous avons fixé aléatoirement les bornes des intervalles en déplaçant les bornes vers la gauche par d'une distance de $u\varepsilon$ où u provient d'une loi uniforme sur l'intervalle (0,1). Notons finalement que la valeur n'est pas toujours la même, mais qu'elle varie peu entre les simulations.

3.2.1.1. *Fonction de perte 0-1 (L_1)*

Les résultats que nous obtenons pour la fonction de perte 0-1 (L_1), qui sont aussi valides pour la fonction de perte 0-k (L_2), concordent avec nos résultats théoriques du chapitre 2. En effet, peu importe le nombre d'intervalles (qui est inversement proportionnel à la taille de la partition ε), nous avons que l'estimateur ponctuel (le mode) est dans l'intervalle du test d'hypothèses multiples 10 000 fois sur 10 000 pour la distribution unimodale symétrique.

Comme prévu, nous n'obtenons pas le même résultat pour les deux autres types de distribution. Notons aussi qu'étant donné la façon dont les simulations ont été programmées, les résultats pour le type bimodal (BI) indiquent le nombre de fois où un des modes est dans l'intervalle, pas nécessairement le mode global.

3.2.1.2. *Fonctions de perte L_3 et L_4*

Ces fonctions de perte dépendant essentiellement des constantes choisies pour chaque intervalle, il n'y a rien à illustrer avec des simulations et nous n'en présentons donc aucune.

3.2.1.3. *Fonction de perte quadratique (L_5)*

Les résultats sont exactement ce à quoi nous nous attendions. L'espérance (*a posteriori*) est toujours dans l'intervalle retenu, peu importe la distribution et le nombre d'intervalles.

3.2.1.4. *Fonction de perte LINEX (L_6)*

Nous voyons que pour les trois types de distribution, le nombre de fois où l'intervalle retenu contient l'estimateur ponctuel augmente lorsque la constante α diminue ou le nombre d'intervalles augmente (ce qui implique que la taille de la partition ε diminue). Nous explorons davantage ceci dans la section suivante.

Tableau 3.2.1: Nombre de fois où l'estimateur ponctuel est dans l'intervalle retenu par le test à hypothèses multiples, pour les trois types de distributions, sur 10 000 simulations.

Nbre intervalles	6	10	20	50	100	250
Dist. US	$\theta X \sim N(\bar{X}, 1)$					
L1	10000	10000	10000	10000	10000	10000
L5	10000	10000	10000	10000	10000	10000
L6, $\alpha = 0.5$	9360	9613	9810	9937	9965	9983
L6, $\alpha = 1$	8792	9290	9601	9822	9926	9961
L6, $\alpha = 2$	7708	8549	9241	9691	9848	9929
Dist. UA	$\theta X \sim G(500 * 20 + \sqrt{500}, 500 * \bar{X} + 1)$					
L1	9976	9989	9987	9998	9998	10000
L5	10000	10000	10000	10000	10000	10000
L6, $\alpha = 0.5$	9793	9861	9944	9979	9986	9994
L6, $\alpha = 1$	9583	9751	9861	9938	9974	9988
L6, $\alpha = 2$	9192	9470	9735	9903	9946	9975
Dist. BI	$\theta X \sim 0,8 * N(\frac{9*\bar{X}}{9+7.2}, 4) + 0,2 * N(5, 4 + 0,1 * \bar{X}, 0,9)$					
L1	8955	9639	9713	9639	9528	9323
L5	10000	10000	10000	10000	10000	10000
L6, $\alpha = 0.5$	8527	9058	9526	9806	9899	9961
L6, $\alpha = 1$	7165	8182	9066	9613	9809	9916
L6, $\alpha = 2$	5328	6728	8242	9243	9631	9840

3.2.2. Simulation pour la fonction de perte LINEX (L_6)

Les résultats sont présentés au tableau 3.2.2 et servent à illustrer les résultats que nous avons obtenus à la proposition 3.1.4 sur la probabilité que l'intervalle retenu par le test d'hypothèses multiples contienne l'estimateur ponctuel pour la fonction de perte LINEX.

Pour les distributions représentant chaque type de distribution (elles sont différentes de celles utilisées pour la section précédente), nous présentons le nombre de fois (sur 10 000) où il y avait accord entre les deux méthodes, le nombre prévu par l'équation de la proposition 3.1.4 et les bornes d'un intervalle de confiance à 95% en utilisant une approximation normale. Si nous notons p la probabilité obtenue avec l'équation de la proposition 3.1.4, nous avons que l'intervalle de confiance à 95% est

$$\left(10000p - z_{0,975}[10000 \times p \times (1 - p)]^{\frac{1}{2}}; 10000p + z_{0,975}[10000 \times p \times (1 - p)]^{\frac{1}{2}} \right).$$

La constante α de la LINEX est fixée à 2 (pour obtenir de plus grandes valeurs de $\alpha\varepsilon$ et la valeur de ε est inversement proportionnelle au nombre d'intervalles. Si I est le nombre d'intervalles,

$$\varepsilon = \begin{cases} \frac{10}{I} & \text{pour le type US;} \\ \frac{18}{I} & \text{pour le type UA;} \\ \frac{21}{I} & \text{pour le type BI.} \end{cases}$$

Notons que ceci fait que les intervalles sont plus larges pour la distribution bimodale que pour la distribution unimodale asymétrique, qui sont eux-mêmes plus larges que ceux de la distribution unimodale symétrique, pour un même nombre d'intervalles.

Dans 17 cas sur 18, l'exception étant la distribution unimodale symétrique avec 250 intervalles, le nombre de fois où il y avait accord entre les deux méthodes est dans l'intervalle de confiance. Dans le dernier cas (les nombres dans les boîtes

grises), la borne supérieure est dépassée de justesse (1 cas). Les résultats des simulations concordent avec le résultat de la proposition 3.1.4.

Tableau 3.2.2: Nombre de fois où l'estimateur ponctuel de la fonction de perte LINEX est dans l'intervalle retenu par le test à hypothèses multiples, pour les trois types de distribution, en utilisant $\alpha = 2$.

Nbre d'intervalles	5	10	20	50	100	250
Dist. US	N(0, 1)					
Nbre obtenu	7055	8409	9177	9666	9849	9951
Nbre prévu	7024	8386	9174	9667	9833	9933
Borne inf. 97.5%	6932	8312	9118	9631	9808	9917
Borne sup. 97.5%	7115	8459	9229	9703	9859	9950
Dist. UA	Gamma(15, 2)					
Nbre obtenu	5479	7294	8568	9419	9706	9882
Nbre prévu	5486	7270	8539	9403	9700	9880
Borne inf. 97.5%	5386	7181	8468	9355	9666	9858
Borne sup. 97.5%	5585	7359	8609	9450	9734	9902
Dist. BI	0,65*N(8, 4) + 0,35*N(14, 1)					
Nbre obtenu	5122	6894	8363	9312	9623	9855
Nbre prévu	5068	6906	8310	9304	9651	9860
Borne inf. 97.5%	4968	6813	8235	9253	9614	9837
Borne sup. 97.5%	5168	6998	8385	9355	9687	9884

3.2.3. Illustration des résultats pour les trois types de distribution

Les tableaux 3.2.3, 3.2.4 et 3.2.5 présentent les pertes espérées (ou les constantes équivalentes pour L_3 et L_4) pour des distributions normale (US), gamma (UA) et un mélange de deux normales (BI) respectivement, lorsque nous utilisons 20 intervalles.

Afin de ne pas surcharger les pages de chiffres, nous avons séparé les tableaux en deux parties en répétant les deux premières colonnes, la première donnant les bornes inférieures des intervalles (l'intervalle s'étendant donc de cette borne jusqu'à la borne indiquée à la ligne suivante, qui est la borne inférieure de l'intervalle suivant) et la deuxième donnant la probabilité *a posteriori* de chaque intervalle. Comme on peut le voir, presque toute la probabilité *a posteriori* est contenue dans ces intervalles.

Pour la première des deux parties de chaque tableau, les troisième et quatrième colonnes présentent les constantes spécifiques à chaque intervalle tel que la perte espérée *a posteriori* de cet intervalle soit égale à celle de l'intervalle ayant la probabilité *a posteriori* maximale auquel nous avons assigné une constante de 1. Pour L_3 (la troisième colonne), si la constante k_i (voir équation (2.1.1)) de l'intervalle est inférieure à ce nombre, cet intervalle sera préféré à celui ayant la probabilité *a posteriori* maximale. Si la constante est plus grande, c'est l'inverse. Par exemple, si nous regardons l'intervalle $(-3, 79; -3, 29)$ du tableau 3.2.3 et que nous fixons la constante de l'intervalle $(-1, 29; -0, 79)$ à 1, nous préférons l'intervalle $(-3, 79; -3, 29)$ si sa constante est 0,81, mais pas si elle est 0,82.

Pour L_4 , si la constante d'un intervalle est supérieure à la constante indiquée (voir équation (2.1.2)), cet intervalle sera préféré à l'intervalle ayant la probabilité *a posteriori* maximale. Si la constante est inférieure, c'est l'inverse. Finalement, la cinquième et dernière colonne montre la perte espérée *a posteriori* pour L_5 pour chaque intervalle. Notons pour terminer que les constantes pour L_4 ont été limitées à 1 000 000 afin d'éviter les nombres astronomiques. Par exemple,

si nous prenons les mêmes intervalles qu'au paragraphe précédent et que nous fixons la constante de l'intervalle $(-1, 29; -0, 79)$ à 1, nous préférons l'intervalle $(-3, 79; -3, 29)$ si sa constante est 35, mais pas si elle est 34.

Dans la deuxième partie, les trois dernières colonnes donnent les pertes espérées pour chacun des intervalles pour la fonction de perte L_6 avec trois valeurs différentes pour les constantes α : $\frac{1}{2}$, 1, 2.

En dernier lieu, mentionnons que les estimateurs ponctuels correspondants aux fonctions de perte sont donnés au bas des colonnes lorsqu'ils existent.

Les figures 3.2.1 (US, 50 intervalles), 3.2.2 (US, 100 intervalles), 3.2.3 (UA, 50 intervalles), 3.2.4 (UA, 100 intervalles), 3.2.5 (BI, 50 intervalles) et 3.2.6 (BI, 100 intervalles) présentent la probabilité et les pertes espérées *a posteriori* pour chaque intervalle. En fait, pour L_3 , ce sont les constantes calculées comme pour les tableaux précédents. Pour L_4 et L_5 , nous avons rééchelonné les pertes espérées *a posteriori* (la constante dans le cas de L_4) avec un facteur multiplicatif pour les mettre sur une échelle de 0 à 1. Pour L_6 , nous présentons le logarithme naturel de la perte espérée *a posteriori*.

Nous pouvons voir que les estimateurs ponctuels (représentés par des "X") sont toujours très proches des intervalles ayant la perte espérée *a posteriori* minimale. À titre de référence, nous avons indiqué la position du mode des distributions par une ligne droite sur l'abscisse.

Tableau 3.2.3: Illustration des pertes espérées pour les 6 fonctions de perte, avec 500 données simulées provenant d'une normale(0, 500) et 20 intervalles.

Début du tableau 3.2.3 : la probabilité, la perte espérée pour L_5 ainsi que les constantes pour les fonctions de perte L_3 et L_4 .

Bornes inf. des intervalles	Probabilité a <i>posteriori</i> (1 - L1)	L3	L4	L5
(-6.29 ; -5.79]	0.00000	0.8064	$5.587 \cdot 10^5$	14.049
(-5.79 ; -5.29]	0.00000	0.8064	$4.927 \cdot 10^4$	11.575
(-5.29 ; -4.79]	0.00003	0.8064	$5.558 \cdot 10^3$	9.350
(-4.79 ; -4.29]	0.00024	0.8066	$8.015 \cdot 10^2$	7.375
(-4.29 ; -3.79]	0.00131	0.8075	$1.478 \cdot 10^2$	5.650
(-3.79 ; -3.29]	0.00556	0.8109	$3.483 \cdot 10^1$	4.175
(-3.29 ; -2.79]	0.01846	0.8216	$1.049 \cdot 10^1$	2.950
(-2.79 ; -2.29]	0.04797	0.8470	4.036	1.976
(-2.29 ; -1.79]	0.09757	0.8936	1.984	1.251
(-1.79 ; -1.29]	0.15534	0.9547	1.246	0.776
(-1.29 ; -0.79]	0.19360	1.0000	1.000	0.551
(-0.79 ; -0.29]	0.18888	0.9942	1.025	0.576
(-0.29 ; 0.21]	0.14426	0.9423	1.342	0.852
(0.21 ; 0.71]	0.08625	0.8825	2.245	1.377
(0.71 ; 1.21]	0.04036	0.8403	4.796	2.152
(1.21 ; 1.71]	0.01478	0.8185	$1.310 \cdot 10^1$	3.177
(1.71 ; 2.21]	0.00424	0.8098	$4.568 \cdot 10^1$	4.452
(2.21 ; 2.71]	0.00095	0.8072	$2.037 \cdot 10^2$	5.977
(2.71 ; 3.21]	0.00017	0.8065	$1.161 \cdot 10^3$	7.753
(3.21 ; 3.71]	0.00002	0.8064	$8.457 \cdot 10^3$	9.778
	Mode a <i>posteriori</i>			Espérance a <i>posteriori</i>
	-0.8404			-0.8404

Fin du tableau 3.2.3 : la probabilité ainsi que la perte espérée pour la fonction de perte LINEX pour trois valeurs de α différentes.

Bornes inf. des intervalles	Probabilité a <i>posteriori</i> (1 - L1)	L6	L6	L6
		$\alpha = 0.5$	$\alpha = 1$	$\alpha = 2$
(-6.29 ; -5.79]	0.00000	0.842	2.104	4.700
(-5.79 ; -5.29]	0.00000	0.729	1.857	4.200
(-5.29 ; -4.79]	0.00003	0.619	1.612	3.701
(-4.79 ; -4.29]	0.00024	0.514	1.370	3.202
(-4.29 ; -3.79]	0.00131	0.415	1.134	2.706
(-3.79 ; -3.29]	0.00556	0.322	0.906	2.217
(-3.29 ; -2.79]	0.01846	0.239	0.692	1.747
(-2.79 ; -2.29]	0.04797	0.168	0.502	1.328
(-2.29 ; -1.79]	0.09757	0.112	1.049	1.049
(-1.79 ; -1.29]	0.15534	0.075	0.264	1.150
(-1.29 ; -0.79]	0.19360	0.064	0.282	2.282
(-0.79 ; -0.29]	0.18888	0.085	0.475	6.221
(-0.29 ; 0.21]	0.14426	0.147	0.954	$1.778 \cdot 10^1$
(0.21 ; 0.71]	0.08625	0.263	1.907	$5.008 \cdot 10^1$
(0.71 ; 1.21]	0.04036	0.447	3.641	$1.387 \cdot 10^2$
(1.21 ; 1.71]	0.01478	0.719	6.661	$3.805 \cdot 10^2$
(1.71 ; 2.21]	0.00424	1.104	11.803	$1.039 \cdot 10^3$
(2.21 ; 2.71]	0.00095	1.633	20.442	$2.829 \cdot 10^3$
(2.71 ; 3.21]	0.00017	2.349	34.848	$7.695 \cdot 10^3$
(3.21 ; 3.71]	0.00002	3.302	58.762	$2.093 \cdot 10^4$
	Mode a <i>posteriori</i>	Estimateurs LINEX		
		$\alpha = 0.5$	$\alpha = 1$	$\alpha = 2$
	-0.8404	-1.0904	-1.3404	-1.8404

Tableau 3.2.4: Illustration des pertes espérées pour les 6 fonctions de perte, avec 500 données simulées provenant d'une $\text{gamma}(20, \theta)$ et 20 intervalles.

Début du tableau 3.2.4 : la probabilité, la perte espérée pour L_5 ainsi que les constantes pour les fonctions de perte L_3 et L_4 .

Bornes inf. des intervalles	Probabilité a <i>posteriori</i> (1 - L1)	L3	L4	L5
(26.74 ; 26.97]	0.00000	0.6961	$> 1.000 \cdot 10^6$	1.3757
(26.97 ; 27.21]	0.00000	0.6961	$> 1.000 \cdot 10^6$	1.1255
(27.21 ; 27.44]	0.00000	0.6961	$> 1.000 \cdot 10^6$	0.9009
(27.44 ; 27.67]	0.00000	0.6961	$> 1.000 \cdot 10^6$	0.7018
(27.67 ; 27.91]	0.00000	0.6961	$2.348 \cdot 10^5$	0.5283
(27.91 ; 28.14]	0.00005	0.6962	$5.846 \cdot 10^3$	0.3803
(28.14 ; 28.38]	0.00107	0.6969	$2.847 \cdot 10^2$	0.2580
(28.38 ; 28.61]	0.01136	0.7041	$2.674 \cdot 10^1$	0.1612
(28.61 ; 28.84]	0.06358	0.7434	4.779	0.0900
(28.84 ; 29.08]	0.18943	0.8588	1.604	0.0443
(29.08 ; 29.31]	0.30386	1.0000	1.000	0.0242
(29.31 ; 29.55]	0.26505	0.9472	1.146	0.0297
(29.55 ; 29.78]	0.12678	0.7972	2.397	0.0607
(29.78 ; 30.01]	0.03349	0.7203	9.074	0.1174
(30.01 ; 30.25]	0.00492	0.6996	$6.181 \cdot 10^1$	0.1996
(30.25 ; 30.48]	0.00040	0.6964	$7.534 \cdot 10^2$	0.3073
(30.48 ; 30.71]	0.00002	0.6962	$1.634 \cdot 10^4$	0.4406
(30.71 ; 30.95]	0.00000	0.6961	$6.268 \cdot 10^5$	0.5995
(30.95 ; 31.18]	0.00000	0.6961	$> 1.000 \cdot 10^6$	0.7840
(31.18 ; 31.42]	0.00000	0.6961	$> 1.000 \cdot 10^6$	0.9940
	Mode a <i>posteriori</i>			Espérance a <i>posteriori</i>
	29.2582			29.2611

Fin du tableau 3.2.4 : la probabilité ainsi que la perte espérée pour la fonction de perte LINEX pour trois valeurs de α différentes.

Bornes inf. des intervalles	Probabilité a <i>posteriori</i> (1 - L1)	L6	L6	L6
		$\alpha = 0.5$	$\alpha = 1$	$\alpha = 2$
(26.74 ; 26.97]	0.00000	0.118	0.351	0.893
(26.97 ; 27.21]	0.00000	0.100	0.302	0.785
(27.21 ; 27.44]	0.00000	0.082	0.254	0.678
(27.44 ; 27.67]	0.00000	0.066	0.209	0.572
(27.67 ; 27.91]	0.00000	0.051	0.166	0.468
(27.91 ; 28.14]	0.00005	0.038	0.126	0.368
(28.14 ; 28.38]	0.00107	0.027	0.090	0.273
(28.38 ; 28.61]	0.01136	0.017	0.059	0.186
(28.61 ; 28.84]	0.06358	0.010	0.034	0.112
(28.84 ; 29.08]	0.18943	0.005	0.018	0.060
(29.08 ; 29.31]	0.30386	0.003	0.011	0.042
(29.31 ; 29.55]	0.26505	0.004	0.016	0.079
(29.55 ; 29.78]	0.12678	0.008	0.038	0.202
(29.78 ; 30.01]	0.03349	0.017	0.079	0.465
(30.01 ; 30.25]	0.00492	0.030	0.146	0.949
(30.25 ; 30.48]	0.00040	0.048	0.245	1.787
(30.48 ; 30.71]	0.00002	0.071	0.384	3.190
(30.71 ; 30.95]	0.00000	0.101	0.575	5.495
(30.95 ; 31.18]	0.00000	0.138	0.830	9.240
(31.18 ; 31.42]	0.00000	0.183	1.166	15.282
	Mode a <i>posteriori</i>	Estimateurs LINEX		
		$\alpha = 0.5$	$\alpha = 1$	$\alpha = 2$
	29.2582	29.2398	29.2185	29.1760

Tableau 3.2.5: Illustration des pertes espérées pour les 6 fonctions de perte, avec 500 données simulées provenant d'un mélange de deux normales et 20 intervalles.

Début du tableau 3.2.5 : la probabilité, la perte espérée pour L_5 ainsi que les constantes pour les fonctions de perte L_3 et L_4 .

Bornes inf. des intervalles	Probabilité a <i>posteriori</i> (1 - L1)	L3	L4	L5
(-10.16 ; -9.15]	0.00000	0.8441	$9.174 \cdot 10^4$	120.431
(-9.15 ; -8.14]	0.00002	0.8441	$9.505 \cdot 10^3$	99.548
(-8.14 ; -7.14]	0.00012	0.8442	$1.265 \cdot 10^3$	80.720
(-7.14 ; -6.13]	0.00072	0.8447	$2.162 \cdot 10^2$	63.949
(-6.13 ; -5.12]	0.00329	0.8469	$4.744 \cdot 10^1$	49.233
(-5.12 ; -4.11]	0.01167	0.8541	$1.336 \cdot 10^1$	36.573
(-4.11 ; -3.10]	0.03226	0.8723	4.831	25.969
(-3.10 ; -2.09]	0.06953	0.9072	2.242	17.421
(-2.09 ; -1.08]	0.11677	0.9557	1.335	10.929
(-1.08 ; -0.07]	0.15283	0.9964	1.020	6.493
(-0.07 ; 0.94]	0.15588	1.0000	1.000	4.113
(0.94 ; 1.95]	0.12394	0.9635	1.258	3.789
(1.95 ; 2.96]	0.07761	0.9151	2.009	5.520
(2.96 ; 3.97]	0.04797	0.8866	3.250	9.308
(3.97 ; 4.98]	0.06382	0.9017	2.443	15.151
(4.98 ; 5.98]	0.08509	0.9226	1.832	23.050
(5.98 ; 6.99]	0.04791	0.8866	3.254	33.006
(6.99 ; 8.00]	0.00983	0.8525	$1.585 \cdot 10^1$	45.017
(8.00 ; 9.01]	0.00072	0.8447	$2.168 \cdot 10^2$	59.084
(9.01 ; 10.02]	0.00002	0.8441	$7.907 \cdot 10^3$	75.207
	Mode a <i>posteriori</i>			Espérance a <i>posteriori</i>
	0.0092			1.0975

Fin du tableau 3.2.5 : la probabilité ainsi que la perte espérée pour la fonction de perte LINEX pour trois valeurs de α différentes.

Bornes inf. des intervalles	Probabilité a <i>posteriori</i> (1 - L1)	L6	L6	L6
		$\alpha = 0.5$	$\alpha = 1$	$\alpha = 2$
(-10.16 ; -9.15]	0.00000	4.429	9.847	$2.070 \cdot 10^1$
(-9.15 ; -8.14]	0.00002	3.927	8.829	$1.867 \cdot 10^1$
(-8.14 ; -7.14]	0.00012	3.430	7.812	$1.663 \cdot 10^1$
(-7.14 ; -6.13]	0.00072	2.940	6.799	$1.460 \cdot 10^1$
(-6.13 ; -5.12]	0.00329	2.463	5.794	$1.259 \cdot 10^1$
(-5.12 ; -4.11]	0.01167	2.007	4.815	$1.079 \cdot 10^1$
(-4.11 ; -3.10]	0.03226	1.586	3.903	$1.055 \cdot 10^1$
(-3.10 ; -2.09]	0.06953	1.224	3.177	$2.199 \cdot 10^1$
(-2.09 ; -1.08]	0.11677	0.957	2.962	$1.214 \cdot 10^2$
(-1.08 ; -0.07]	0.15283	0.850	4.147	$8.831 \cdot 10^2$
(-0.07 ; 0.94]	0.15588	1.008	9.175	$6.629 \cdot 10^3$
(0.94 ; 1.95]	0.12394	1.602	$2.474 \cdot 10^1$	$4.989 \cdot 10^4$
(1.95 ; 2.96]	0.07761	2.922	$6.923 \cdot 10^1$	$3.755 \cdot 10^5$
(2.96 ; 3.97]	0.04797	5.441	$1.931 \cdot 10^2$	$2.827 \cdot 10^6$
(3.97 ; 4.98]	0.06382	9.948	$5.346 \cdot 10^2$	$2.128 \cdot 10^7$
(4.98 ; 5.98]	0.08509	$1.775 \cdot 10^1$	$1.473 \cdot 10^3$	$1.601 \cdot 10^8$
(5.98 ; 6.99]	0.04791	$3.100 \cdot 10^1$	$4.050 \cdot 10^3$	$1.205 \cdot 10^9$
(6.99 ; 8.00]	0.00983	$5.329 \cdot 10^1$	$1.112 \cdot 10^4$	$9.072 \cdot 10^9$
(8.00 ; 9.01]	0.00072	$9.054 \cdot 10^1$	$3.053 \cdot 10^4$	$6.828 \cdot 10^{10}$
(9.01 ; 10.02]	0.00002	$1.526 \cdot 10^2$	$8.376 \cdot 10^4$	$5.140 \cdot 10^{11}$
	Mode a <i>posteriori</i>	Estimateurs LINEX		
		$\alpha = 0.5$	$\alpha = 1$	$\alpha = 2$
	0.0092	-0.5667	-1.7679	-3.8792

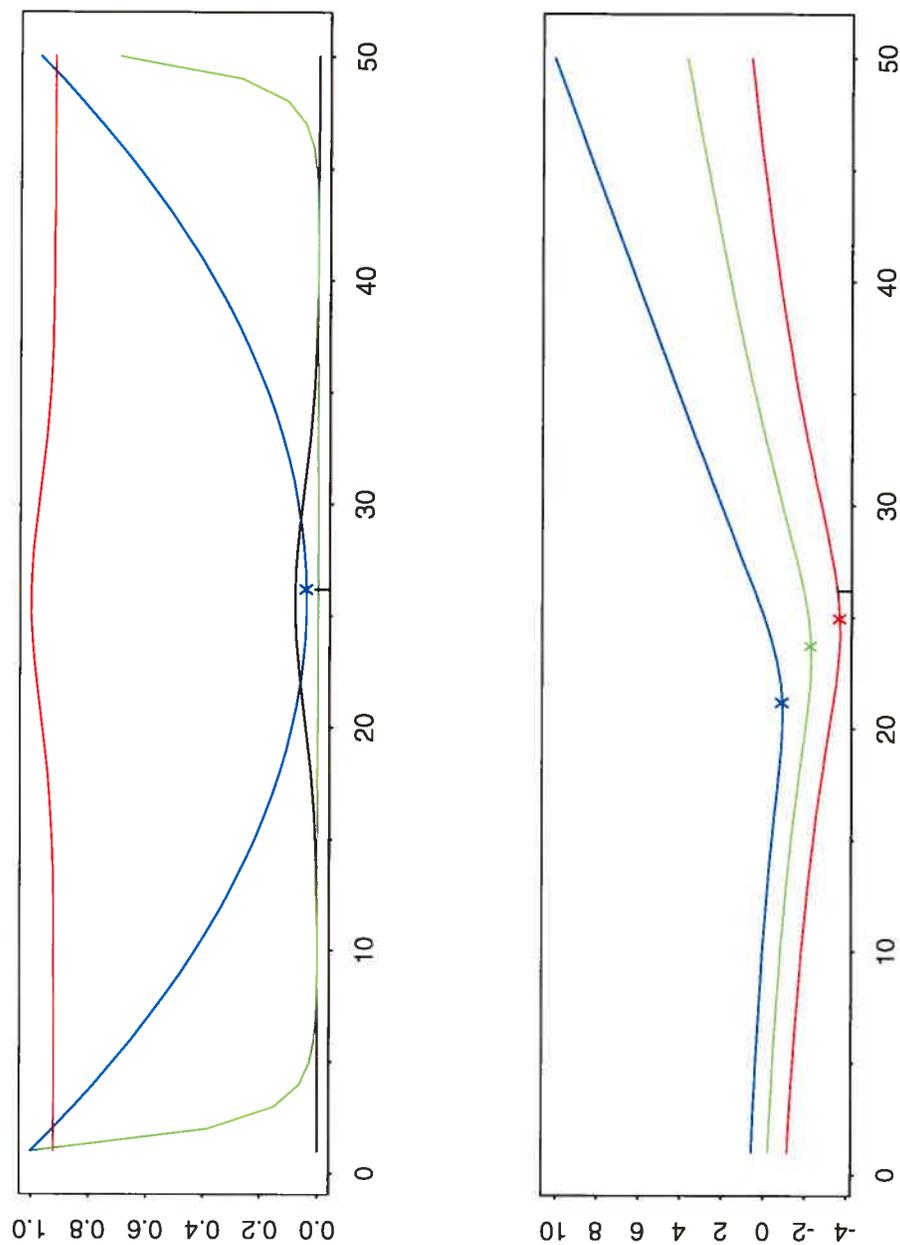


FIGURE 3.2.1. Probabilité *a posteriori* (gauche, en noir), constantes pour L_3 (gauche, en rouge) et L_4 (gauche, en vert), pertes espérées *a posteriori* pour L_5 (à gauche, en bleu), L_6 (à droite, $\alpha = 0.5$ en rouge, $\alpha = 1$ en vert, $\alpha = 2$ en bleu) pour chacun des 50 intervalles, pour la distribution unimodale symétrique

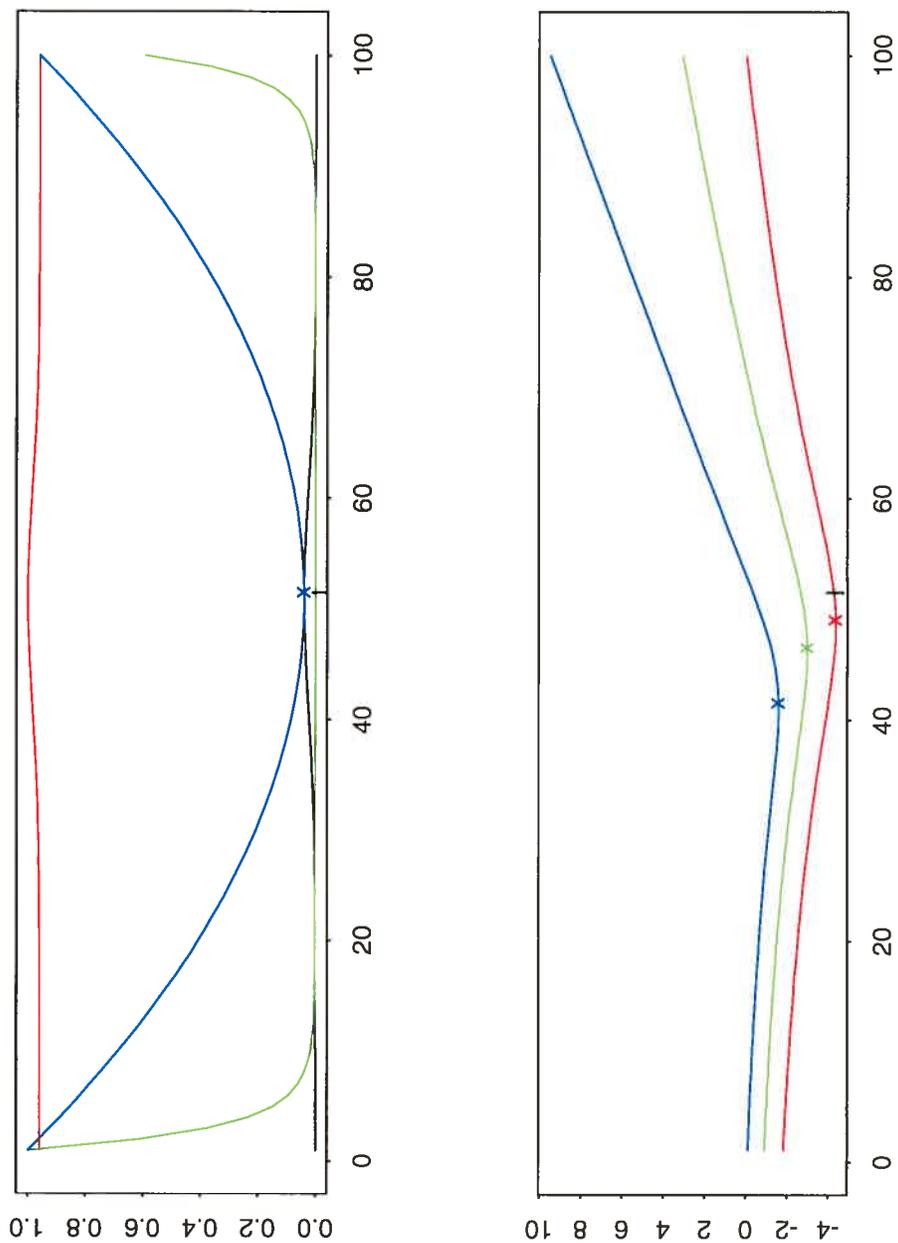


FIGURE 3.2.2. Probabilité *a posteriori* (gauche, en noir), constantes pour L_3 (gauche, en rouge) et L_4 (gauche, en vert), pertes espérées *a posteriori* pour L_5 (à gauche, en bleu), L_6 (à droite, $\alpha = 0.5$ en rouge, $\alpha = 1$ en vert, $\alpha = 2$ en bleu) pour chacun des 100 intervalles, pour la distribution unimodale symétrique

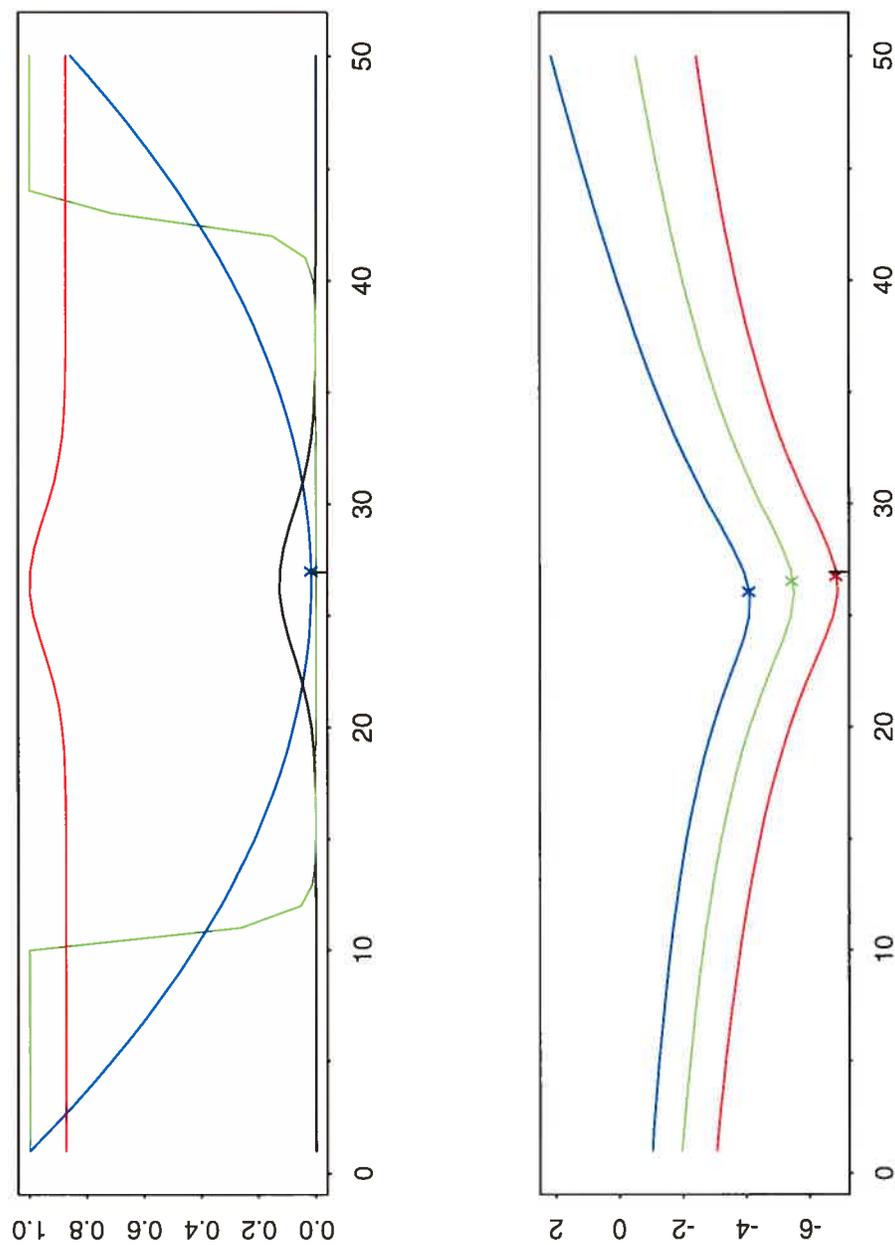


FIGURE 3.2.3. Probabilité *a posteriori* (gauche, en noir), constantes pour L_3 (gauche, en rouge) et L_4 (gauche, en vert), pertes espérées *a posteriori* pour L_5 (à gauche, en bleu), L_6 (à droite, $\alpha = 0.5$ en rouge, $\alpha = 1$ en vert, $\alpha = 2$ en bleu) pour chacun des 50 intervalles, pour la distribution unimodale asymétrique

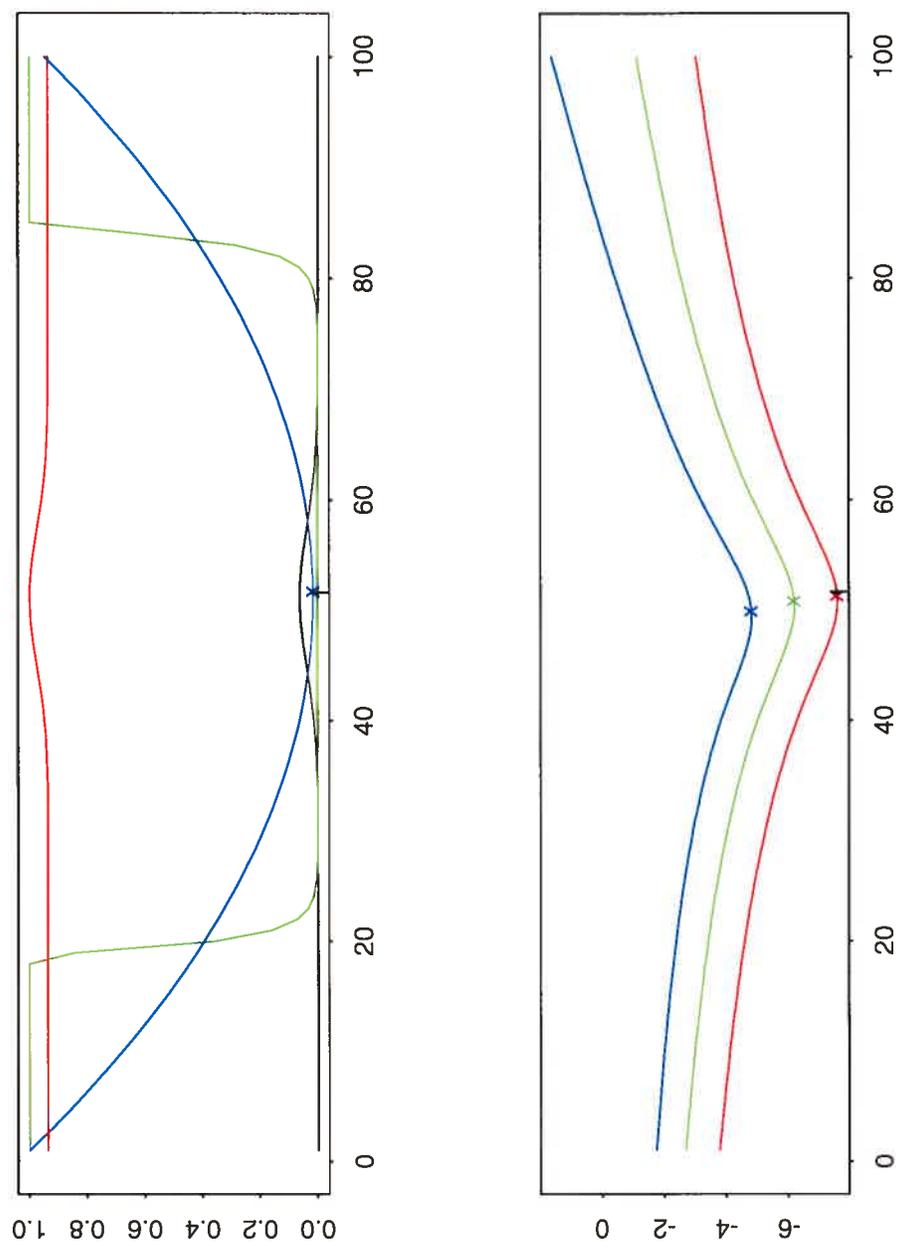


FIGURE 3.2.4. Probabilité *a posteriori* (gauche, en noir), constantes pour L_3 (gauche, en rouge) et L_4 (gauche, en vert), pertes espérées *a posteriori* pour L_5 (à gauche, en bleu), L_6 (à droite, $\alpha = 0.5$ en rouge, $\alpha = 1$ en vert, $\alpha = 2$ en bleu) pour chacun des 100 intervalles, pour la distribution unimodale asymétrique

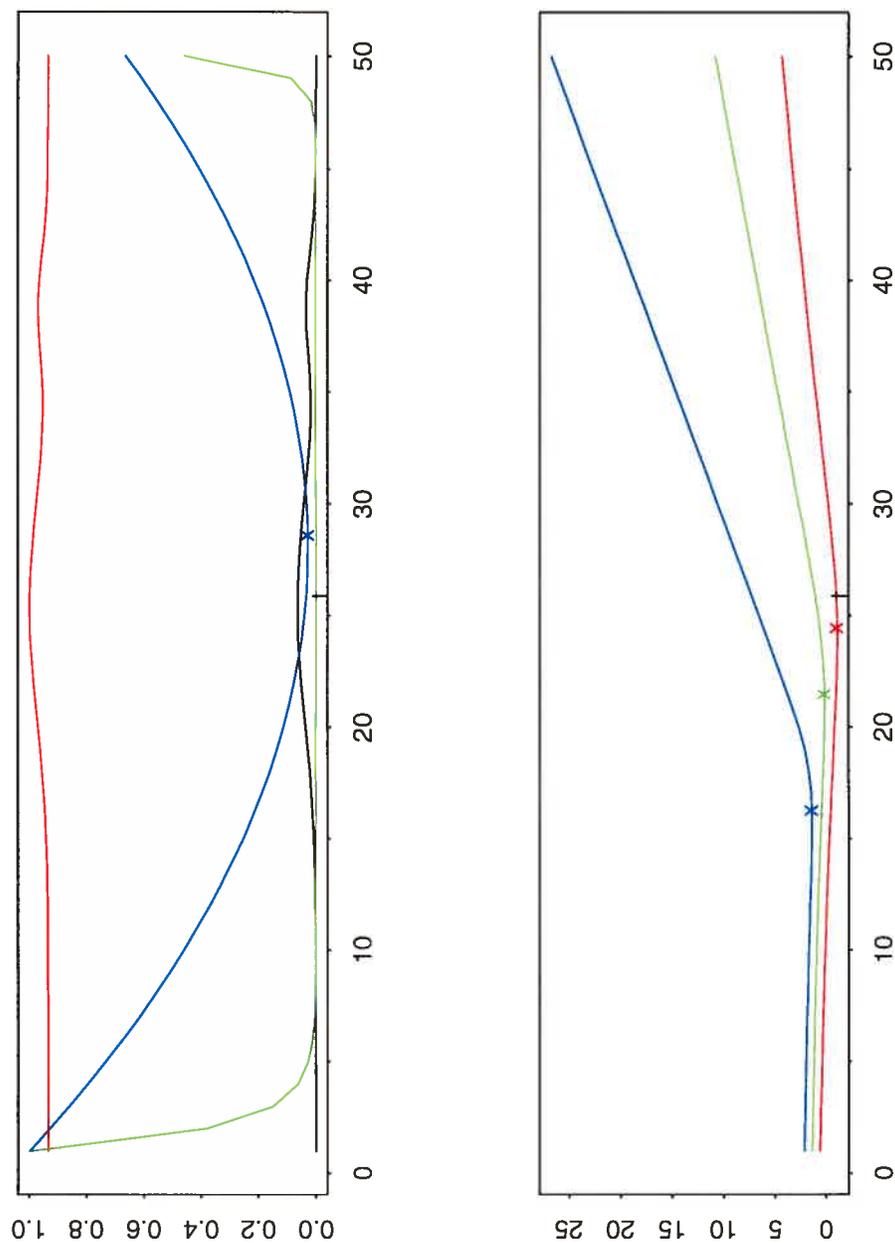


FIGURE 3.2.5. Probabilité *a posteriori* (gauche, en noir), constantes pour L_3 (gauche, en rouge) et L_4 (gauche, en vert), pertes espérées *a posteriori* pour L_5 (à gauche, en bleu). L_6 (à droite, $\alpha = 0.5$ en rouge, $\alpha = 1$ en vert, $\alpha = 2$ en bleu) pour chacun des 50 intervalles, pour la distribution bimodale

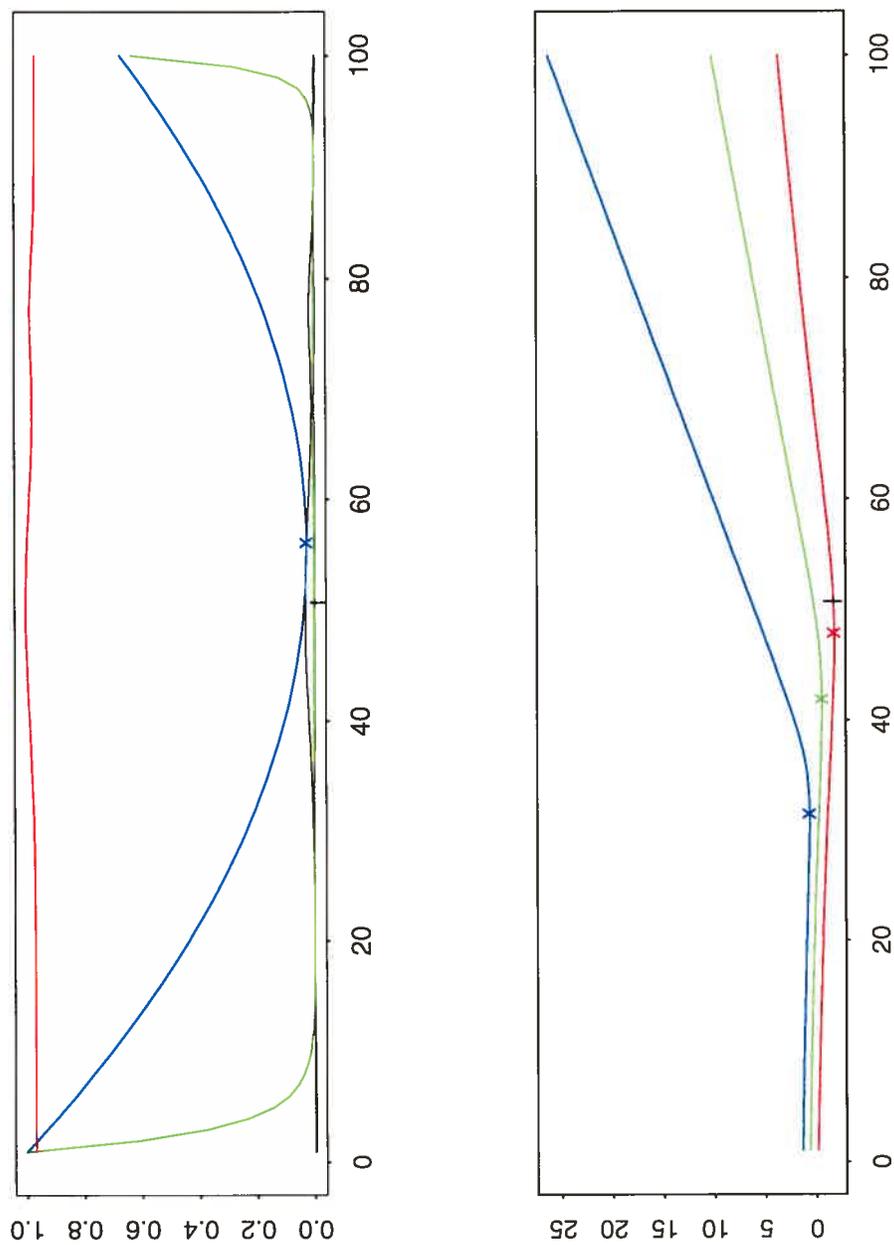


FIGURE 3.2.6. Probabilité *a posteriori* (gauche, en noir), constantes pour L_3 (gauche, en rouge) et L_4 (gauche, en vert), pertes espérées *a posteriori* pour L_5 (à gauche, en bleu), L_6 (à droite, $\alpha = 0.5$ en rouge, $\alpha = 1$ en vert, $\alpha = 2$ en bleu) pour chacun des 100 intervalles, pour la distribution bimodale

3.3. EXEMPLE

Nous avons appliqué nos résultats à un jeu de données réelles, provenant d'une ligne d'assurance-incendie d'une compagnie d'assurance danoise, tirées d'un article de Ramlau-Hansen (1988). Nous avons décidé d'utiliser une distribution gamma, qui reste une distribution simple et dont l'ajustement aux données est acceptable (voir figure 3.3.1).

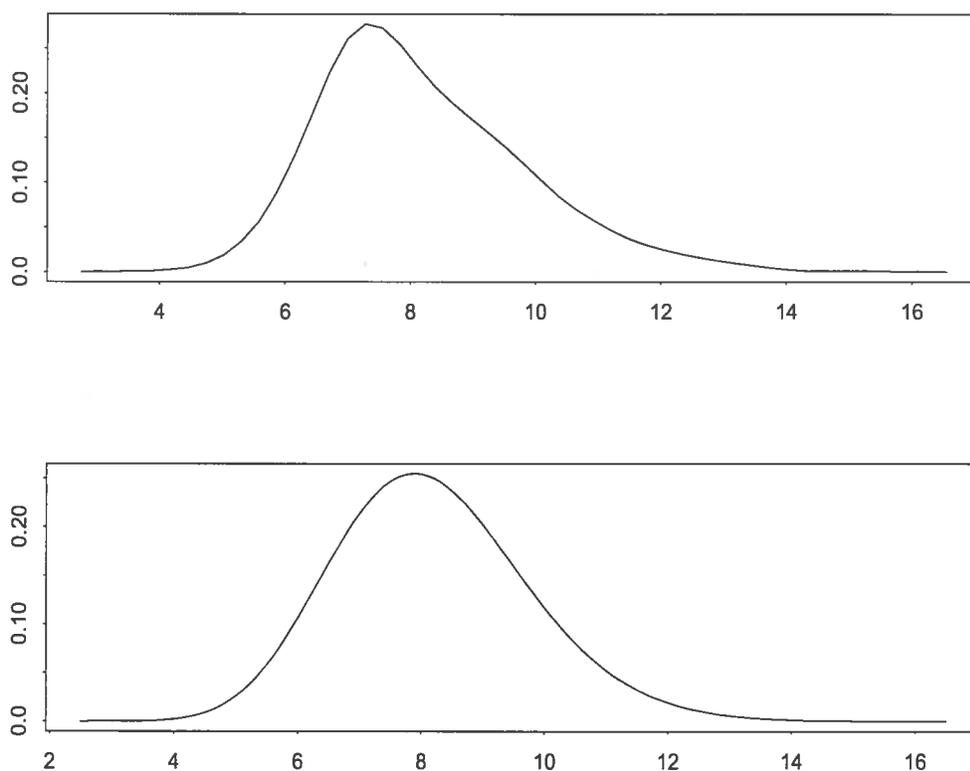


FIGURE 3.3.1. Comparaison entre la distribution du logarithme des données réelles et le modèle. Le graphique du haut présente la distribution du logarithme des données (ajusté par la fonction `plot(density)` de S-plus, qui utilise un noyau gaussien avec une fenêtre calculée selon la règle de Silverman) et celui du bas présente la densité du modèle ajusté aux données.

Nous retrouvons les résultats au tableau 3.3.1, qui est construit de la même façon que les tableaux de la section 3.2.3.

Tableau 3.3.1: Illustration des pertes espérées pour les 6 fonctions de perte, avec le jeu de données réelles et 20 intervalles.

Début du tableau 3.3.1 : la probabilité, la perte espérée pour L_5 ainsi que les constantes pour les fonctions de perte L_3 et L_4 .

Bornes inf. des intervalles	Probabilité a <i>posteriori</i> (1 - L1)	L3	L4	L5
(3.14 ; 3.15]	0.00002	0.8256	10502.111	$9.27 \cdot 10^{-5}$
(3.15 ; 3.16]	0.00011	0.8256	1655.720	$7.45 \cdot 10^{-5}$
(3.16 ; 3.17]	0.00054	0.8260	321.490	$5.83 \cdot 10^{-5}$
(3.17 ; 3.18]	0.00227	0.8274	76.775	$4.40 \cdot 10^{-5}$
(3.18 ; 3.19]	0.00775	0.8320	22.519	$3.18 \cdot 10^{-5}$
(3.19 ; 3.20]	0.02153	0.8437	8.102	$2.16 \cdot 10^{-5}$
(3.20 ; 3.21]	0.04885	0.8680	3.571	$1.33 \cdot 10^{-5}$
(3.21 ; 3.22]	0.09059	0.9078	1.926	$7.10 \cdot 10^{-6}$
(3.22 ; 3.23]	0.13749	0.9572	1.269	$2.87 \cdot 10^{-6}$
(3.23 ; 3.24]	0.17099	0.9958	1.020	$6.36 \cdot 10^{-7}$
(3.24 ; 3.25]	0.17444	1.0000	1.000	$4.04 \cdot 10^{-7}$
(3.25 ; 3.26]	0.14618	0.9669	1.193	$2.17 \cdot 10^{-6}$
(3.26 ; 3.27]	0.10072	0.9180	1.732	$5.94 \cdot 10^{-6}$
(3.27 ; 3.28]	0.05713	0.8756	3.053	$1.17 \cdot 10^{-5}$
(3.28 ; 3.29]	0.02671	0.8482	6.531	$1.95 \cdot 10^{-5}$
(3.29 ; 3.30]	0.01030	0.8341	16.934	$2.92 \cdot 10^{-5}$
(3.30 ; 3.31]	0.00328	0.8283	53.156	$4.10 \cdot 10^{-5}$
(3.31 ; 3.32]	0.00086	0.8263	201.812	$5.48 \cdot 10^{-5}$
(3.32 ; 3.33]	0.00019	0.8257	925.730	$7.05 \cdot 10^{-5}$
(3.33 ; 3.34]	0.00003	0.8256	5125.235	$8.83 \cdot 10^{-5}$
	Mode a <i>posteriori</i>			Espérance a <i>posteriori</i>
	3.2410			3.2412

Fin du tableau 3.3.1 : la probabilité ainsi que la perte espérée pour la fonction de perte LINEX pour trois valeurs de α différentes.

Bornes inf. des intervalles	Probabilité a <i>posteriori</i> (1 - L1)	L6	L6	L6
		$\alpha = 0.5$	$\alpha = 1$	$\alpha = 2$
(3.14 ; 3.15]	0.00002	$1.20 \cdot 10^{-5}$	$4.71 \cdot 10^{-5}$	$1.82 \cdot 10^{-4}$
(3.15 ; 3.16]	0.00011	$9.76 \cdot 10^{-6}$	$3.84 \cdot 10^{-5}$	$1.49 \cdot 10^{-4}$
(3.16 ; 3.17]	0.00054	$7.77 \cdot 10^{-6}$	$3.06 \cdot 10^{-5}$	$1.19 \cdot 10^{-4}$
(3.17 ; 3.18]	0.00227	$6.03 \cdot 10^{-6}$	$2.38 \cdot 10^{-5}$	$9.27 \cdot 10^{-5}$
(3.18 ; 3.19]	0.00775	$4.52 \cdot 10^{-6}$	$1.79 \cdot 10^{-5}$	$6.99 \cdot 10^{-5}$
(3.19 ; 3.20]	0.02153	$3.26 \cdot 10^{-6}$	$1.29 \cdot 10^{-5}$	$5.06 \cdot 10^{-5}$
(3.20 ; 3.21]	0.04885	$2.25 \cdot 10^{-6}$	$8.90 \cdot 10^{-6}$	$3.50 \cdot 10^{-5}$
(3.21 ; 3.22]	0.09059	$1.48 \cdot 10^{-6}$	$5.86 \cdot 10^{-6}$	$2.31 \cdot 10^{-5}$
(3.22 ; 3.23]	0.13749	$9.54 \cdot 10^{-7}$	$3.79 \cdot 10^{-6}$	$1.50 \cdot 10^{-5}$
(3.23 ; 3.24]	0.17099	$6.79 \cdot 10^{-7}$	$2.71 \cdot 10^{-6}$	$1.08 \cdot 10^{-5}$
(3.24 ; 3.25]	0.17444	$6.53 \cdot 10^{-7}$	$2.62 \cdot 10^{-6}$	$1.05 \cdot 10^{-5}$
(3.25 ; 3.26]	0.14618	$8.78 \cdot 10^{-7}$	$3.53 \cdot 10^{-6}$	$1.43 \cdot 10^{-5}$
(3.26 ; 3.27]	0.10072	$1.35 \cdot 10^{-6}$	$5.46 \cdot 10^{-6}$	$2.22 \cdot 10^{-5}$
(3.27 ; 3.28]	0.05713	$2.08 \cdot 10^{-6}$	$8.41 \cdot 10^{-6}$	$3.43 \cdot 10^{-5}$
(3.28 ; 3.29]	0.02671	$3.07 \cdot 10^{-6}$	$1.24 \cdot 10^{-5}$	$5.07 \cdot 10^{-5}$
(3.29 ; 3.30]	0.01030	$4.31 \cdot 10^{-6}$	$1.74 \cdot 10^{-5}$	$7.14 \cdot 10^{-5}$
(3.30 ; 3.31]	0.00328	$5.80 \cdot 10^{-6}$	$2.35 \cdot 10^{-5}$	$9.66 \cdot 10^{-5}$
(3.31 ; 3.32]	0.00086	$7.56 \cdot 10^{-6}$	$3.07 \cdot 10^{-5}$	$1.26 \cdot 10^{-4}$
(3.32 ; 3.33]	0.00019	$9.57 \cdot 10^{-6}$	$3.89 \cdot 10^{-5}$	$1.61 \cdot 10^{-4}$
(3.33 ; 3.34]	0.00003	$1.18 \cdot 10^{-5}$	$4.82 \cdot 10^{-5}$	$2.00 \cdot 10^{-4}$
	Mode a <i>posteriori</i>	Estimateurs LINEX		
		$\alpha = 0.5$	$\alpha = 1$	$\alpha = 2$
	3.2410	3.2410	3.2409	3.2407

3.4. CONCLUSION

Nous avons obtenu plusieurs résultats intéressants et utiles dans ce chapitre, puis nous les avons illustrés à l'aide de simulations. Nous avons aussi étudié un exemple réel.

Nous savons maintenant que les tests d'hypothèses multiples et l'estimation s'accordent toujours lorsque nous utilisons une fonction de perte quadratique (L_5) ou 0-1 (L_1) (mais uniquement avec une distribution unimodale symétrique). Pour les distributions qui ne sont pas unimodales symétriques et la fonction de perte 0-1 (L_1), ce n'est pas toujours le cas. Malgré tout, grâce au théorème que l'on retrouve en annexe, il y a un mode dans l'intervalle ayant la probabilité *a posteriori* maximale ou un des intervalles adjacents, peu importe la taille de la partition ε .

Pour la fonction de perte LINEX (L_6), nous savons que si l'estimateur ponctuel n'est pas dans l'intervalle ayant la perte espérée *a posteriori* minimale, il sera dans l'intervalle immédiatement à droite (si $\alpha > 0$) ou à gauche (si $\alpha < 0$) et que la probabilité qu'il se trouve dans l'intervalle retenu augmente lorsque la taille de la partition ε décroît.

CONCLUSION

Dans ce mémoire nous nous sommes intéressés aux tests d'hypothèses multiples sur des partitions de l'espace-paramètre en intervalles de largeurs égales. Nous avons aussi comparé les résultats de ces tests avec les estimateurs ponctuels utilisant les fonctions de perte correspondantes.

Au premier chapitre nous avons présenté les 6 fonctions de perte que nous avons utilisées dans le mémoire ainsi que quelques résultats préliminaires sur ces fonctions de perte.

Dans le deuxième chapitre, nous avons appliqué les résultats du chapitre 1 à trois types de distributions (quatre dans le cas de la fonction de perte 0-1) afin d'obtenir des résultats un peu plus spécifiques.

Au dernier chapitre, nous avons comparé les estimateurs ponctuels avec les intervalles retenus par les tests d'hypothèses multiples lorsque la largeur des intervalles tend vers 0. Ensuite, nous avons illustré ces résultats avec des simulations ainsi qu'un exemple basé sur des données réelles.

Finalement, à l'annexe, nous présentons un théorème très général sur le nombre d'intervalles (ainsi que lesquels exactement) pouvant être l'intervalle ayant la probabilité *a posteriori* maximale. Les seules exigences sont que le nombre de modes soit fini, la densité soit continue et qu'elle converge vers 0 aux deux extrémités du support.

Nous avons trouvé les décisions de Bayes pour les tests d'hypothèses multiples pour nos fonctions de perte. Pour les fonctions de perte L_3 et L_4 , ces résultats sont peu spécifiques étant donné que les constantes assignées à chaque intervalle sont très importantes pour le résultat. Nous avons obtenu des résultats encore plus spécifiques pour les fonctions de perte L_1 , L_2 , L_5 et L_6 pour les types de distributions unimodale symétrique (US), unimodale asymétrique (UA) et bimodale (BI). Pour les fonctions de perte L_1 et L_2 , nous avons aussi des résultats pour toute distribution multimodale (MM).

Les simulations que nous avons faites au chapitre 3 confirment nos résultats sur l'accord entre les estimateurs ponctuels et les tests d'hypothèses multiples. Dans le cas de L_1 et L_2 , il y a toujours accord pour les distributions unimodales symétriques, mais pas toujours pour les autres distributions, et il est impossible de prédire s'il y aura accord. Pour L_3 et L_4 , il est impossible de trouver des résultats généraux satisfaisants. Pour L_5 , la fonction de perte quadratique, l'estimateur ponctuel est toujours dans l'intervalle retenu par le test d'hypothèses multiples. Pour L_6 , il n'y a pas toujours accord, mais l'estimateur ponctuel est dans l'intervalle retenu par le test ou un intervalle adjacent dans tous les cas. De plus, nous pouvons donner la probabilité que l'estimateur ponctuel soit dans l'intervalle retenu par le test, et cette probabilité dépend de la constante α de la fonction de perte L_6 et de la largeur des intervalles, que nous avons notée ε .

En conclusion, l'accord entre les estimateurs ponctuels et les tests d'hypothèses multiples dépend de la fonction de perte et parfois de la largeur de la partition. De plus, nous sommes d'accord avec Evans (1997) qu'en général, les estimateurs ponctuels sont moins compliqués que les tests d'hypothèses multiples.

Annexe A

THÉORÈME SUR LES DISTRIBUTIONS MULTIMODALES

Dans cette annexe, nous présentons le théorème qui abolit la restriction sur la « proximité » des modes et la position des minima du théorème de la section 2.4.

Étant donné que l'abolition de cette restriction ajoute beaucoup de complexité, nous présentons de courtes explications et des résumés avant les étapes importantes. Il est tout de même fortement suggéré de lire la section 2.4 comme préparation.

Commençons par expliquer la sélection des β_i . De façon générale, $\beta_i = \alpha_i + 1$ si la densité à la borne de droite de l'intervalle Θ_{α_i} est (strictement) plus élevée que celle à la borne de gauche de cet intervalle. De même, $\beta_i = \alpha_i - 1$ si c'est la densité à la borne de gauche qui est (strictement) plus élevée. Nous établissons la convention que β_i a priorité sur β_{i+1} . Ceci signifie que si $\beta_i = \beta_{i+1}$, nous ne retiendrons pas β_{i+1} . Finalement, puisqu'ils se trouvent aux extrémités (du point de vue des modes), β_1 et β_M ont droit à un traitement spécifique.

Ajoutons que si deux intervalles contenant (au moins) un mode sont adjacents ou identiques, on utilise les β_i pour couvrir les deux intervalles adjacents à gauche et à droite. Procédons maintenant à l'énoncé du théorème.

Théorème A.0.1. Soit $\pi(\theta|x)$ la densité (continue) d'une distribution ayant M modes, notés $\theta_1, \theta_2, \dots, \theta_M$ tel que $\theta_1 < \theta_2 < \dots < \theta_M$.

Définissons un ensemble d'intervalles $\Theta_i, i \in \mathbb{Z}$ qui forment une partition de l'espace-paramètre Θ avec les caractéristiques supplémentaires suivantes :

- $\Theta_i = (l_i, u_i], \quad u_i - l_i = \varepsilon, \quad \forall i \in \mathbb{Z};$
- $\Theta_{i+1} = (l_{i+1} = l_i + \varepsilon = u_i, u_{i+1} = u_i + \varepsilon], \quad \forall i \in \mathbb{Z}.$

Ainsi, il existe M constantes $\alpha_1 \leq \alpha_2 \leq \dots \leq \alpha_M \in \mathbb{Z}$ tel que $\theta_i \in \Theta_{\alpha_i}, i = 1, 2, \dots, M$.

Pour chaque α_i , il nous est possible de choisir un seul β_i tel que nous limitons $\arg \max_{a \in \mathbb{Z}} \Pr\{\theta \in \Theta_a|x\}$ à l'ensemble $\{\alpha_1, \alpha_2, \dots, \alpha_M, \beta_1, \beta_2, \dots, \beta_M\}$ qui ne comporte que $2M$ éléments (au maximum).

Nous définissons les β_i de la manière suivante :

$$\beta_1 = \begin{cases} \alpha_1 - 1 & \text{si } \alpha_2 - \alpha_1 \leq 1; \\ \left\{ \begin{array}{l} \alpha_1 - 1 & \text{si } \pi(l_{\alpha_1}|x) > \pi(u_{\alpha_1}|x); \\ \alpha_1 + 1 & \text{si } \pi(l_{\alpha_1}|x) \leq \pi(u_{\alpha_1}|x); \end{array} \right\} & \text{si } \alpha_2 - \alpha_1 \geq 2. \end{cases}$$

$$\beta_i = \begin{cases} \text{pas défini} & \text{si } \alpha_{i+1} - \alpha_i \leq 1 \text{ et } \alpha_i - \alpha_{i-1} \leq 1; \\ \alpha_i + 1 & \text{si } \alpha_{i+1} - \alpha_i \geq 2 \text{ et } \alpha_i - \alpha_{i-1} \leq 1; \\ \alpha_i - 1 & \text{si } \alpha_{i+1} - \alpha_i \leq 1 \text{ et } \alpha_i - \alpha_{i-1} \geq 2 \text{ et } \beta_{i-1} \neq \alpha_i - 1; \\ \left\{ \begin{array}{l} \alpha_i - 1 & \text{si } \pi(l_{\alpha_i}|x) > \pi(u_{\alpha_i}|x) \text{ et } \beta_{i-1} \neq \alpha_i - 1; \\ \alpha_i + 1 & \text{si } \pi(l_{\alpha_i}|x) < \pi(u_{\alpha_i}|x); \end{array} \right\} & \text{si } \alpha_{i+1} - \alpha_i \geq 2 \\ & \text{et } \alpha_i - \alpha_{i-1} \geq 2; \\ \text{pas défini} & \text{sinon.} \end{cases}$$

$$\beta_M = \left\{ \begin{array}{l} \alpha_M + 1 \text{ si } \alpha_M - \alpha_{M-1} \leq 1; \\ \left. \begin{array}{l} \left. \begin{array}{l} \alpha_M - 1 \\ \alpha_M + 1 \end{array} \right\} \text{ si } \pi(l_{\alpha_M}|x) \geq \pi(u_{\alpha_M}|x) \\ \text{et } \beta_{M-1} \neq \alpha_M - 1; \\ \alpha_M + 1 \text{ si } \pi(l_{\alpha_M}|x) < \pi(u_{\alpha_M}|x); \\ \text{pas défini sinon.} \end{array} \right\} \text{ si } \alpha_M - \alpha_{M-1} \geq 2.$$

Avant d'amorcer la preuve, nous présentons un aperçu de la façon dont nous allons procéder.

Nous commençons par montrer que l'intervalle de probabilité *a posteriori* maximale ne peut être à gauche (ou avoir un indice plus petit, ce qui est équivalent) des intervalles que nous avons choisis. Nous séparons selon la valeur que prend β_1 .

Ensuite, nous démontrons que l'intervalle de probabilité *a posteriori* maximale ne peut être à droite (avoir un indice plus élevé) des intervalles que nous avons choisis. Nous séparons aussi selon la valeur de β_M . La preuve est très similaire à celle de la partie précédente, mais pas tout à fait parce que β_M n'est pas défini exactement symétriquement à β_1 .

Finalement, nous nous attaquons aux intervalles non choisis entre les paires de modes adjacents. Nous montrons que dans tous les cas (ou presque), leur probabilité *a posteriori* est inférieure à au moins une des probabilités *a posteriori* des intervalles choisis qui les bornent. Les quelques exceptions sont des cas très particuliers qui sont explorés et couverts dans le lemme qui suit. Il nous permettra ensuite d'alléger la preuve par quelques hypothèses ou présupposés qui, s'ils ne sont pas justifiés, renvoient au lemme. Ces hypothèses ou présupposés seront notés explicitement. Voici donc le lemme et sa preuve, et la preuve du théorème suivra.

Lemme A.0.1. *Soit la même situation que dans l'énoncé du théorème (incluant les définitions des β_i).*

Si, pour un certain intervalle Θ_{α_I} , donc ayant pour indice α_I ,

- $\alpha_{I+1} - \alpha_I \geq 2$ et $\alpha_I - \alpha_{I-1} \geq 2$;
- $\min_{\theta \in (\theta_I, u_{\alpha_I})} \pi(\theta|x) = \pi(u_{\alpha_I}|x)$, c'est-à-dire que le premier minimum à droite de θ_I n'est pas dans Θ_{α_I} (ou que ce minimum est un plateau);
- $\beta_I \neq \alpha_I + 1$, ce qui implique $\pi(l_{\alpha_I}|x) \geq \pi(u_{\alpha_I}|x)$,

alors nous avons que $\max_{a \in \{\alpha_1, \alpha_2, \dots, \alpha_I, \beta_1, \beta_2, \dots, \beta_I\}} \Pr\{\theta \in \Theta_a|x\} \geq \varepsilon \pi(u_{\alpha_I}|x)$.

Démonstration (du lemme) :

Puisque $\pi(l_{\alpha_I}|x) \geq \pi(u_{\alpha_I}|x)$ et que $\min_{\theta \in (\theta_I, u_{\alpha_I})} \pi(\theta|x) = \pi(u_{\alpha_I}|x)$, nous avons l'implication suivante :

$$\begin{aligned} \Pr\{\theta \in \Theta_{\alpha_I}|x\} &< \varepsilon \pi(u_{\alpha_I}|x) \\ \Rightarrow \min_{\theta \in (l_{\alpha_I}, \theta_I)} \pi(\theta|x) &< \pi(l_{\alpha_I}|x). \end{aligned}$$

En effet,

$$\begin{aligned} \Pr\{\theta \in \Theta_{\alpha_I}|x\} &= \int_{\Theta_{\alpha_I}} \pi(\theta|x) d\theta \\ &= \int_{l_{\alpha_I}}^{\theta_I} \pi(\theta|x) d\theta + \int_{\theta_I}^{u_{\alpha_I}} \pi(\theta|x) d\theta \\ &\geq \int_{l_{\alpha_I}}^{\theta_I} \pi(\theta|x) d\theta + (u_{\alpha_I} - \theta_I) \pi(u_{\alpha_I}|x) \end{aligned}$$

et donc

$$\begin{aligned} \Pr\{\theta \in \Theta_{\alpha_I}|x\} &< \varepsilon \pi(u_{\alpha_I}|x) \\ \Rightarrow \int_{l_{\alpha_I}}^{\theta_I} \pi(\theta|x) d\theta &< (\theta_I - l_{\alpha_I}) \pi(u_{\alpha_I}|x) \\ \Rightarrow \min_{\theta \in (l_{\alpha_I}, \theta_I)} \pi(\theta|x) &< \pi(u_{\alpha_I}|x) \leq \pi(l_{\alpha_I}|x) \end{aligned}$$

car $\varepsilon = u_{\alpha_I} - l_{\alpha_I} = (u_{\alpha_I} - \theta_I) + (\theta_I - l_{\alpha_I})$.

Tout ceci nous indique que le minimum immédiatement à gauche de θ_I (notons le θ^*) est dans l'intervalle Θ_{α_I} , ce qui implique que

$$\pi(\theta_{I-1}|x) \geq \pi(s|x) \geq \pi(t|x) \geq \pi(\theta^*|x), \theta_{I-1} \leq s \leq t \leq \theta^*.$$

Si $\beta_I = \alpha_I - 1$ (donc $\pi(l_{\alpha_I}|x) > \pi(u_{\alpha_I}|x)$), nous avons $\Pr\{\theta \in \Theta_{\beta_I}|x\} \geq \varepsilon\pi(l_{\alpha_I}|x) > \varepsilon\pi(u_{\alpha_I}|x)$ puisqu'il n'y a aucun minimum entre θ_{I-1} et celui situé dans Θ_{α_I} , ce qui implique que $\pi(\theta|x)$ est décroissante dans Θ_{β_I} . Donc, si $\beta_I = \alpha_I - 1$, le lemme tient ;

Si par contre $\beta_I \neq \alpha_I - 1$, nous avons que soit $\beta_{I-1} = \alpha_I - 1$ (ce qui implique $\beta_{I-1} = \alpha_{I-1} + 1$), soit $\pi(l_{\alpha_I}|x) = \pi(u_{\alpha_I}|x)$, ou les deux. Étant donné la complexité du deuxième cas, nous le présentons en dernier. Nous commençons par le cas où les deux conditions sont réalisées, puis par le premier cas.

– Si $\beta_{I-1} = \alpha_{I-1} + 1$ et que $\pi(l_{\alpha_I}|x) = \pi(u_{\alpha_I}|x)$, nous avons que

$$\Pr\{\theta \in \Theta_{\beta_{I-1}}|x\} \geq \varepsilon\pi(u_{\beta_{I-1}}|x) \geq \varepsilon\pi(l_{\alpha_I}|x) = \varepsilon\pi(u_{\alpha_I}|x).$$

car $\pi(\theta|x)$ est décroissante sur l'intervalle (θ_{I-1}, θ^*) et $(l_{\beta_{I-1}}, l_{\alpha_I}) \subset (\theta_{I-1}, \theta^*)$;

– Si $\beta_{I-1} = \alpha_{I-1} + 1$ et que $\pi(l_{\alpha_I}|x) > \pi(u_{\alpha_I}|x)$, alors le signe d'égalité à la fin de l'équation précédente est remplacé par un signe d'inégalité ">" ;

– Sinon, $\beta_{I-1} \neq \alpha_{I-1} + 1$ (ce qui implique $\pi(l_{\alpha_{I-1}}|x) \geq \pi(u_{\alpha_{I-1}}|x)$ et $\pi(l_{\alpha_I}|x) = \pi(u_{\alpha_I}|x)$), et si nous avons que

$$\Pr\{\theta \in \Theta_{\alpha_{I-1}}|x\} \geq \varepsilon\pi(u_{\alpha_{I-1}}|x) \geq \varepsilon\pi(l_{\alpha_I}|x) = \varepsilon\pi(u_{\alpha_I}|x).$$

alors le lemme est vérifié. Si ce n'est pas le cas, nous réunissons à nouveau les conditions du lemme, mais pour α_{I-1} au lieu de α_I . Nous pouvons continuer ainsi jusqu'à α_1 et β_1 , qui est défini différemment des autres β_i .

Alors, si $\beta_1 = \alpha_1 + 1$, nous avons que

$$\Pr\{\theta \in \Theta_{\beta_1}|x\} \geq \varepsilon\pi(u_{\beta_1}|x) \geq \varepsilon\pi(l_{\alpha_1}|x) = \varepsilon\pi(u_{\alpha_1}|x)$$

et sinon, $\beta_1 = \alpha_1 - 1$, ce qui implique que $\pi(l_{\alpha_1}|x) > \pi(u_{\alpha_1}|x)$, et donc

$$\Pr\{\theta \in \Theta_{\alpha_1}|x\} \geq \varepsilon\pi(u_{\alpha_1}|x) \geq \varepsilon\pi(l_{\alpha_1}|x) = \varepsilon\pi(u_{\alpha_1}|x).$$

Il est impossible que si $\pi(l_{\alpha_1}|x) > \pi(u_{\alpha_1}|x)$ et le minimum entre θ_1 et θ_2 n'est pas dans l'intervalle Θ_{α_1} , nous ayons

$$\Pr\{\theta \in \Theta_{\alpha_1}|x\} < \varepsilon\pi(u_{\alpha_1}|x)$$

puisqu'il n'y a pas de mode θ^* tel que $\theta^* < \theta_1$.

Le lemme tient aussi dans l'autre sens, en passant de α_I à α_{I+1} et ainsi de suite jusqu'à α_M et β_M .

Démonstration (du théorème) :

Comme nous l'avons déjà mentionné, nous commençons par le côté gauche de la distribution.

Puisqu'il n'y a pas de mode à gauche de θ_1 , nous savons que

$$\pi(s|x) \leq \pi(t|x) \leq \pi(\theta_1|x), s \leq t \leq \theta_1.$$

Séparons maintenant selon que $\beta_1 = \alpha_1 - 1$ ou $\beta_1 = \alpha_1 + 1$.

Premièrement, si $\beta_1 = \alpha_1 - 1$, pour un intervalle Θ_a « à gauche » de l'intervalle Θ_{β_1} , c'est-à-dire $a < \beta_1$, nous aurons que

$$\Pr\{\theta \in \Theta_a|x\} \leq \varepsilon\pi(u_a|x) \leq \varepsilon\pi(l_{\beta_1}|x) \leq \Pr\{\theta \in \Theta_{\beta_1}|x\}$$

puisque $\Theta_a \subset \{\theta : \theta < \Theta_1\} \supset \Theta_{\beta_1}$.

Deuxièmement, si $\beta_1 = \alpha_1 + 1$, nous savons que $\alpha_2 - \alpha_1 \geq 2$ et que $\pi(u_{\alpha_1}|x) \geq \pi(l_{\alpha_1}|x)$. Nous avons que

$$\Pr\{\theta \in \Theta_a|x\} \leq \varepsilon\pi(u_a|x) \leq \varepsilon\pi(l_{\alpha_1}|x) \quad \forall a < \alpha_1.$$

Nous avons maintenant deux possibilités :

$$\varepsilon\pi(l_{\alpha_1}|x) \leq \Pr\{\theta \in \Theta_{\alpha_1}|x\}$$

et alors

$$\Pr\{\theta \in \Theta_a|x\} \leq \Pr\{\theta \in \Theta_{\alpha_1}|x\} \quad \forall a < \alpha_1;$$

$$\varepsilon\pi(l_{\alpha_1}|x) > \Pr\{\theta \in \Theta_{\alpha_1}|x\}$$

$$\Rightarrow \min_{\theta \in (\theta_1, u_{\alpha_1})} \pi(\theta|x) < \pi(l_{\alpha_1}|x) \leq \pi(u_{\alpha_1}|x)$$

puisque nous savons que $\min_{\theta \in (l_{\alpha_1}, \theta_1)} \pi(\theta|x) = \pi(l_{\alpha_1}|x)$. Ceci veut dire que la distribution atteint son minimum local entre θ_1 et u_{α_1} . Mais, puisque $\theta_2 > u_{\beta_1}$, nous aurons que

$$\varepsilon\pi(l_{\alpha_1}|x) \leq \varepsilon\pi(u_{\alpha_1}|x) = \varepsilon\pi(l_{\beta_1}|x) \leq \Pr\{\theta \in \Theta_{\beta_1}|x\}$$

et donc

$$\Pr\{\theta \in \Theta_a|x\} \leq \Pr\{\theta \in \Theta_{\beta_1}|x\} \quad \forall a < \alpha_1.$$

Nous avons montré que $\arg \max_{a \in \mathbb{Z}} \Pr\{\theta \in \Theta_a|x\} \notin \min\{\alpha_1, \beta_1\}$.

Nous passons maintenant au côté droit de la distribution.

Étant donné que θ_M est le mode le plus à droite, nous avons que

$$\pi(\theta_M|x) \geq \pi(t|x) \geq \pi(s|x), \theta_M \geq t \geq s.$$

Nous séparons maintenant selon que $\beta_M = \alpha_M + 1$ ou non.

En premier lieu, si $\beta_M = \alpha_M + 1$,

$$\Pr\{\theta \in \Theta_a|x\} \leq \varepsilon\pi(l_a|x) \leq \varepsilon\pi(u_{\beta_M}|x) \leq \Pr\{\theta \in \Theta_{\beta_M}|x\} \quad \forall a > \beta_M$$

puisque $\Theta_a \subset \{\theta : \theta > \theta_M\} \supset \Theta_{\beta_M}$.

En deuxième lieu, nous avons que soit $\beta_M = \alpha_M - 1$, soit $\beta_{M-1} = \alpha_M - 1$. Nous savons alors que $\alpha_M - \alpha_{M-1} \geq 2$ et que $\pi(l_{\alpha_M}|x) \geq \pi(u_{\alpha_M}|x)$.

Similairement au côté gauche, nous avons que

$$\Pr\{\theta \in \Theta_a|x\} \leq \varepsilon\pi(l_a|x) \leq \varepsilon\pi(u_{\alpha_M}|x) \quad \forall a > \alpha_M.$$

Encore une fois, deux possibilités s'offrent à nous :

– soit

$$\varepsilon\pi(u_{\alpha_M}|x) \leq \Pr\{\theta \in \Theta_{\alpha_M}|x\}$$

et donc

$$\Pr\{\theta \in \Theta_a|x\} \leq \Pr\{\theta \in \Theta_{\alpha_M}|x\} \quad \forall a > \alpha_M;$$

– ou encore, soit

$$\varepsilon\pi(u_{\alpha_M}|x) > \Pr\{\theta \in \Theta_{\alpha_M}|x\}.$$

Nous tirons une conclusion similaire à celle du côté gauche, c'est-à-dire que nous avons atteint le minimum local entre l_{α_M} et θ_M .

De même, puisque $\theta_{M-1} > \alpha_M - 1$, nous aurons que

$$\varepsilon\pi(u_{\alpha_M}|x) \leq \varepsilon\pi(l_{\alpha_M}|x) = \varepsilon\pi(u_{\alpha_{M-1}}|x) \leq \Pr\{\theta \in \Theta_{\alpha_{M-1}}|x\}$$

et donc

$$\Pr\{\theta \in \Theta_a|x\} \leq \Pr\{\theta \in \Theta_{\alpha_{M-1}}|x\} \quad \forall a > \alpha_M.$$

Or, nous savons que soit $\beta_M = \alpha_M - 1$, soit $\beta_{M-1} = \alpha_M - 1$, ce qui veut dire que nous venons de montrer que $\arg \max_{a \in Z} \Pr\{\theta \in \Theta_a|x\} \neq \max\{\alpha_M, \beta_M\}$.

Nous arrivons maintenant à la troisième et dernière partie de la preuve. Nous montrons maintenant que pour α_i et α_{i+1} quelconques (et β_i et β_{i+1} définis comme

dans l'énoncé), nous avons que

$$\arg \max_{a \in \mathbb{Z}} \Pr\{\theta \in \Theta_a | x\} \notin \{\max\{\alpha_i, \beta_i\} + 1, \max\{\alpha_i, \beta_i\} + 2, \dots, \max\{\alpha_{i+1}, \beta_{i+1}\} - 1\}.$$

Nous divisons en plusieurs parties selon le nombre d'intervalles entre Θ_{α_i} et $\Theta_{\alpha_{i+1}}$ (ou, plus précisément, selon $\alpha_{i+1} - \alpha_i$). Selon les cas, nous divisons encore selon les valeurs de β_i et β_{i+1} ainsi que la position du « minimum ».

Premièrement, si $\alpha_{i+1} - \alpha_i \leq 1$, il n'y a aucun intervalle entre Θ_{α_i} et $\Theta_{\alpha_{i+1}}$ puisqu'ils sont soit adjacents soit superposés.

Deuxièmement, si $\alpha_{i+1} - \alpha_i = 2$, alors il y a un seul intervalle entre les deux. Si $\beta_i = \alpha_i + 1$ ou $\beta_{i+1} = \alpha_{i+1} - 1$, cet unique intervalle aura été choisi et nous n'avons rien à démontrer. Nous supposons donc que $\beta_i \neq \alpha_i + 1$ et $\beta_{i+1} \neq \alpha_{i+1} - 1$. Ceci implique plusieurs choses :

- $\alpha_i - \alpha_{i-1} \geq 2$ (si $i \neq 1$) et $\alpha_{i+2} - \alpha_{i+1} \geq 2$ (si $i \neq M - 1$);
- $\pi(l_{\alpha_i} | x) \geq \pi(u_{\alpha_i} | x)$ et $\pi(l_{\alpha_{i+1}} | x) \leq \pi(u_{\alpha_{i+1}} | x)$.

Puisqu'il n'y a pas de mode entre θ_i et θ_{i+1} , $\exists \theta^* \in (\theta_i, \theta_{i+1})$ tel que

$$\pi(\theta_i | x) \geq \pi(s | x) \geq \pi(t | x) \geq \pi(\theta^* | x), \theta_i \leq s \leq t \leq \theta^*$$

et

$$\pi(\theta^* | x) \leq \pi(s | x) \leq \pi(t | x) \leq \pi(\theta_{i+1} | x), \theta^* \leq s \leq t \leq \theta_{i+1}.$$

Il y a trois possibilités pour la position de θ^* . Si $\theta^* \in \Theta_{\alpha_i}$, nous avons que

$$\begin{aligned} \Pr\{\theta \in \Theta_{\alpha_{i+1}} | x\} &\leq \varepsilon \pi(u_{\alpha_{i+1}} | x) \\ &= \varepsilon \pi(u_{\alpha_{i+1}-1} | x) \\ &= \varepsilon \pi(l_{\alpha_{i+1}} | x) \\ &\leq \Pr\{\theta \in \Theta_{\alpha_{i+1}} | x\}. \end{aligned}$$

Cette dernière inégalité peut ne pas être respectée, mais nous nous retrouvons alors avec les conditions du lemme présenté précédemment, qui implique que $\Pr\{\theta \in \Theta_{\alpha_{i+1}} | x\} \leq \max_{a \in \{\alpha_{i+1}, \alpha_{i+2}, \dots, \alpha_M, \beta_{i+1}, \beta_{i+2}, \dots, \beta_M\}} \Pr\{\theta \in \Theta_a | x\}$. Par souci de

simplicité, nous maintenons ces abus de notation en mentionnant l'existence du lemme à chaque fois.

Si $\theta^* \in \Theta_{\alpha_{i+1}}$, nous avons de façon similaire que

$$\Pr\{\theta \in \Theta_{\alpha_{i+1}}|x\} \leq \varepsilon \pi(l_{\alpha_{i+1}}|x) = \varepsilon \pi(u_{\alpha_i}|x) \leq \Pr\{\theta \in \Theta_{\alpha_i}|x\} \text{ (lemme)} .$$

Sinon, $\theta^* \notin \Theta_{\alpha_i}$ et $\theta^* \notin \Theta_{\alpha_{i+1}}$, nous avons alors que

$$\pi(\theta|x) \leq \max\{\pi(u_{\alpha_i}|x); \pi(l_{\alpha_{i+1}}|x)\} \forall \theta \in (u_{\alpha_i}, l_{\alpha_{i+1}}).$$

Ceci implique que

$$\begin{aligned} \Pr\{\theta \in \Theta_{\alpha_{i+1}}|x\} &\leq \varepsilon \max\{\pi(u_{\alpha_i}|x); \pi(l_{\alpha_{i+1}}|x)\} \\ &\leq \varepsilon \max_{a \in \{\alpha_i, \alpha_{i+1}\}} \Pr\{\theta \in \Theta_a|x\} \text{ (lemme)} \end{aligned}$$

puisque le lemme s'applique à

$$\varepsilon \pi(u_{\alpha_i}|x) \leq \Pr\{\theta \in \Theta_{\alpha_i}|x\} \text{ (lemme)}$$

et

$$\varepsilon \pi(l_{\alpha_{i+1}}|x) \leq \Pr\{\theta \in \Theta_{\alpha_{i+1}}|x\} \text{ (lemme)} .$$

Donc, pour $\alpha_{i+1} - \alpha_i = 2$, $\arg \max_{a \in Z} \Pr\{\theta \in \Theta_a|x\} \neq \alpha_i + 1$, si $\beta_i \neq \alpha_i + 1$ et $\beta_{i+1} \neq \alpha_{i+1} - 1$.

Troisièmement, pour $\alpha_{i+1} - \alpha_i = 3$, si $\beta_i = \alpha_i + 1$ et $\beta_{i+1} = \alpha_{i+1} - 1$, alors il n'y a pas d'intervalles non choisis. Comme pour le cas $\alpha_{i+1} - \alpha_i = 2$, $\exists \theta^* \in (\theta_i, \theta_{i+1})$ tel que

$$\pi(\theta_i|x) \geq \pi(s|x) \geq \pi(t|x) \geq \pi(\theta^*|x), \theta_i \leq s \leq t \leq \theta^*$$

et

$$\pi(\theta^*|x) \leq \pi(s|x) \leq \pi(t|x) \leq \pi(\theta_{i+1}|x), \theta^* \leq s \leq t \leq \theta_{i+1}.$$

Si $\theta^* \in \Theta_{\alpha_i}$, nous avons deux possibilités :

– si $\beta_{i+1} = \alpha_{i+1} - 1$,

$$\Pr\{\theta \in \Theta_{\alpha_{i+1}}|x\} \leq \varepsilon\pi(u_{\alpha_{i+1}}|x) = \varepsilon\pi(l_{\beta_{i+1}}|x) \leq \Pr\{\theta \in \Theta_{\beta_{i+1}}|x\}.$$

Le lemme n'est pas nécessaire ici ;

– sinon,

$$\begin{aligned} \Pr\{\theta \in \Theta_{\alpha_{i+1}}|x\} &\leq \Pr\{\theta \in \Theta_{\alpha_{i+1}-1}|x\} \\ &\leq \varepsilon\pi(u_{\alpha_{i+1}-1}|x) \\ &= \varepsilon\pi(l_{\alpha_{i+1}}|x) \\ &\leq \Pr\{\theta \in \Theta_{\alpha_{i+1}}|x\} \text{ (lemme)}. \end{aligned}$$

Si $\theta^* \in (u_{\alpha_i}, u_{\alpha_{i+1}})$, nous avons trois possibilités :

– si $\beta_i = \alpha_i + 1$ (qui implique $\beta_{i+1} \neq \alpha_{i+1} - 1$),

$$\begin{aligned} \Pr\{\theta \in \Theta_{\alpha_{i+1}-1}|x\} &\leq \varepsilon\pi(u_{\alpha_{i+1}-1}|x) \\ &= \varepsilon\pi(l_{\alpha_{i+1}}|x) \\ &\leq \Pr\{\theta \in \Theta_{\alpha_{i+1}}|x\} \text{ (lemme)}; \end{aligned}$$

– si $\beta_{i+1} = \alpha_{i+1} - 1$ (qui implique $\beta_i \neq \alpha_i + 1$),

$$\begin{aligned} \Pr\{\theta \in \Theta_{\alpha_{i+1}}|x\} &\leq \varepsilon\pi(l_{\alpha_{i+1}}|x) \\ &= \varepsilon\pi(u_{\alpha_i}|x) \\ &\leq \Pr\{\theta \in \Theta_{\alpha_i}|x\} \text{ (lemme)}; \end{aligned}$$

– si $\beta_i \neq \alpha_i + 1$ et $\beta_{i+1} \neq \alpha_{i+1} - 1$, nous savons que

$$\pi(\theta|x) \leq \max\{\pi(u_{\alpha_i}|x); \pi(l_{\alpha_{i+1}}|x)\}, \quad \forall \theta \in (u_{\alpha_i}, l_{\alpha_{i+1}}).$$

Or, comme

$$\varepsilon\pi(u_{\alpha_i}|x) \leq \Pr\{\theta \in \Theta_{\alpha_i}|x\} \text{ (lemme)}$$

et

$$\varepsilon\pi(l_{\alpha_{i+1}}|x) \leq \Pr\{\theta \in \Theta_{\alpha_{i+1}}|x\} \text{ (lemme)},$$

nous aurons donc que

$$\begin{aligned} \Pr\{\theta \in \Theta_b|x\} &\leq \varepsilon \max\{\pi(u_{\alpha_i}|x); \pi(l_{\alpha_{i+1}}|x)\} \\ &\leq \max_{a \in \{\alpha_i, \alpha_{i+1}\}} \Pr\{\theta \in \Theta_a|x\} \text{ (lemme),} \\ &\text{pour } b = \alpha_i + 1, \alpha_{i+1} - 1. \end{aligned}$$

Si $\theta^* \in (l_{\alpha_{i+1}-1}, l_{\alpha_{i+1}})$, la preuve est très similaire au cas précédent ($\theta^* \in (u_{\alpha_i}, u_{\alpha_{i+1}})$) car nous sommes dans un cas presque symétrique.

Finalement, si $\theta^* \in \Theta_{\alpha_{i+1}}$, la preuve est très similaire au premier cas ($\theta^* \in \Theta_{\alpha_i}$) et nous ne la présentons pas non plus.

Nous avons donc démontré la validité de notre affirmation pour $\alpha_{i+1} - \alpha_i = 3$.

Finalement, nous passons au dernier cas, $\alpha_{i+1} - \alpha_i \geq 4$. Nous avons toujours que $\exists \theta^* \in (\theta_i, \theta_{i+1})$ tel que

$$\pi(\theta_i|x) \geq \pi(s|x) \geq \pi(t|x) \geq \pi(\theta^*|x), \theta_i \leq s \leq t \leq \theta^*$$

et

$$\pi(\theta^*|x) \leq \pi(s|x) \leq \pi(t|x) \leq \pi(\theta_{i+1}|x), \theta^* \leq s \leq t \leq \theta_{i+1}.$$

Nous divisons maintenant selon la position de θ^* . Si $\theta^* \in \Theta_{\alpha_i}$, nous avons deux possibilités :

– si $\beta_{i+1} = \alpha_{i+1} - 1$, nous avons que $\forall a$ tel que $\alpha_i < a < \beta_{i+1}$,

$$\Pr\{\theta \in \Theta_a|x\} \leq \varepsilon \pi(u_a|x) \leq \varepsilon \pi(l_{\beta_{i+1}}|x) \leq \Pr\{\theta \in \Theta_{\beta_{i+1}}|x\};$$

– sinon, nous avons que $\forall \alpha_i < a < \alpha_{i+1}$,

$$\begin{aligned} \Pr\{\theta \in \Theta_{\alpha_a}|x\} &\leq \Pr\{\theta \in \Theta_{\alpha_{i+1}-1}|x\} \\ &\leq \varepsilon \pi(u_{\alpha_{i+1}-1}|x) \\ &= \varepsilon \pi(l_{\alpha_{i+1}}|x) \\ &\leq \Pr\{\theta \in \Theta_{\alpha_{i+1}}|x\} \text{ (lemme).} \end{aligned}$$

Si $\theta^* \in \Theta_{\alpha_i+1} = (u_{\alpha_i}, u_{\alpha_i+1}]$, nous avons quatre possibilités :

- si $\beta_i = \alpha_i + 1$ et $\beta_{i+1} = \alpha_{i+1} - 1$, alors $\forall a$ tel que $\beta_i < a < \beta_{i+1}$, nous avons que

$$\Pr\{\theta \in \Theta_a|x\} \leq \varepsilon\pi(u_a|x) \leq \varepsilon\pi(l_{\beta_{i+1}}|x) \leq \Pr\{\theta \in \Theta_{\beta_{i+1}}|x\};$$

- si $\beta_i = \alpha_i + 1$ et $\beta_{i+1} \neq \alpha_{i+1} - 1$, alors $\forall a$ tel que $\beta_i < a < \alpha_{i+1}$ nous avons que

$$\begin{aligned} \Pr\{\theta \in \Theta_a|x\} &\leq \Pr\{\theta \in \Theta_{\alpha_{i+1}-1}|x\} \\ &\leq \varepsilon\pi(u_{\alpha_{i+1}-1}|x) \\ &= \varepsilon\pi(l_{\alpha_{i+1}}|x) \\ &\leq \Pr\{\theta \in \Theta_{\alpha_{i+1}}|x\} \text{ (lemme);} \end{aligned}$$

- si $\beta_i \neq \alpha_i + 1$ et $\beta_{i+1} = \alpha_{i+1} - 1$, nous savons que

$$\varepsilon\pi(u_{\alpha_i}|x) \leq \Pr\{\theta \in \Theta_{\alpha_i}|x\} \text{ (lemme)}$$

et

$$\varepsilon\pi(l_{\beta_{i+1}}|x) \leq \Pr\{\theta \in \Theta_{\beta_{i+1}}|x\}.$$

Alors, $\forall a$ tel que $\alpha_i < a < \beta_{i+1}$, nous avons que

$$\begin{aligned} \Pr\{\theta \in \Theta_a|x\} &\leq \varepsilon \max\{\pi(u_{\alpha_i}|x); \pi(l_{\beta_{i+1}}|x)\} \\ &\leq \max_{b \in \{\alpha_i, \beta_{i+1}\}} \Pr\{\theta \in \Theta_b|x\} \text{ (lemme);} \end{aligned}$$

- si $\beta_i \neq \alpha_i + 1$ et $\beta_{i+1} \neq \alpha_{i+1} - 1$, nous savons que

$$\varepsilon\pi(u_{\alpha_i}|x) \leq \Pr\{\theta \in \Theta_{\alpha_i}|x\} \text{ (lemme)}$$

et

$$\varepsilon\pi(l_{\alpha_{i+1}}|x) \leq \Pr\{\theta \in \Theta_{\alpha_{i+1}}|x\} \text{ (lemme).}$$

Alors, $\forall a$ tel que $\alpha_i < a < \alpha_{i+1}$, nous avons que

$$\begin{aligned} \Pr\{\theta \in \Theta_a|x\} &\leq \varepsilon \max\{\pi(u_{\alpha_i}|x); \pi(l_{\alpha_{i+1}}|x)\} \\ &\leq \max_{b \in \{\alpha_i, \alpha_{i+1}\}} \Pr\{\theta \in \Theta_b|x\} \text{ (lemme).} \end{aligned}$$

Si $\theta^* \in (u_{\alpha_i+1}, l_{\alpha_i+1-1})$, nous savons que

$$\varepsilon\pi(u_{\alpha_i}|x) \leq \Pr\{\theta \in \Theta_{\alpha_i}|x\} \text{ si } \beta_i \neq \alpha_i + 1 \text{ (lemme),}$$

$$\varepsilon\pi(u_{\alpha_i+1}|x) \leq \Pr\{\theta \in \Theta_{\alpha_i+1}|x\} \text{ si } \beta_i \neq \alpha_i+1 - 1 \text{ (lemme),}$$

$$\varepsilon\pi(u_{\beta_i}|x) \leq \Pr\{\theta \in \Theta_{\beta_i}|x\} \text{ si } \beta_i = \alpha_i + 1$$

et

$$\varepsilon\pi(l_{\beta_i+1}|x) \leq \Pr\{\theta \in \Theta_{\beta_i+1}|x\} \text{ si } \beta_i+1 = \alpha_i+1 - 1.$$

Quatre possibilités se présentent à nous :

– si $\beta_i = \alpha_i + 1$ et $\beta_{i+1} = \alpha_{i+1} - 1$, alors $\forall a$ tel que $\beta_i < a < \beta_{i+1}$, nous avons que

$$\begin{aligned} \Pr\{\theta \in \Theta_a|x\} &\leq \varepsilon \max\{\pi(u_{\beta_i}|x); \pi(l_{\beta_i+1}|x)\} \\ &\leq \max_{b \in \{\beta_i, \beta_i+1\}} \Pr\{\theta \in \Theta_b|x\}; \end{aligned}$$

– si $\beta_i = \alpha_i + 1$ et $\beta_{i+1} \neq \alpha_{i+1} - 1$, alors $\forall a$ tel que $\beta_i < a < \alpha_{i+1}$, nous avons que

$$\begin{aligned} \Pr\{\theta \in \Theta_a|x\} &\leq \varepsilon \max\{\pi(u_{\beta_i}|x); \pi(l_{\alpha_i+1}|x)\} \\ &\leq \max_{b \in \{\beta_i, \alpha_i+1\}} \Pr\{\theta \in \Theta_b|x\} \text{ (lemme);} \end{aligned}$$

– si $\beta_i \neq \alpha_i + 1$ et $\beta_{i+1} = \alpha_{i+1} - 1$, alors $\forall a$ tel que $\alpha_i < a < \beta_{i+1}$, nous avons que

$$\begin{aligned} \Pr\{\theta \in \Theta_a|x\} &\leq \varepsilon \max\{\pi(u_{\alpha_i}|x); \pi(l_{\beta_i+1}|x)\} \\ &\leq \max_{b \in \{\alpha_i, \beta_i+1\}} \Pr\{\theta \in \Theta_b|x\} \text{ (lemme);} \end{aligned}$$

– si $\beta_i \neq \alpha_i + 1$ et $\beta_{i+1} \neq \alpha_{i+1} - 1$, alors $\forall a$ tel que $\alpha_i < a < \alpha_{i+1}$, nous avons que

$$\begin{aligned} \Pr\{\theta \in \Theta_a|x\} &\leq \varepsilon \max\{\pi(u_{\alpha_i}|x); \pi(l_{\alpha_i+1}|x)\} \\ &\leq \max_{b \in \{\alpha_i, \alpha_i+1\}} \Pr\{\theta \in \Theta_b|x\} \text{ (lemme).} \end{aligned}$$

Si $\theta^* \in (l_{\alpha_i+1-1}, l_{\alpha_i+1}] = \Theta_{\alpha_i+1-1}$, nous retournons au 2^e cas, car la preuve est similaire.

Finalement, si $\theta^* \in \Theta_{\alpha_{i+1}}$, la preuve est similaire à celle pour $\theta^* \in \Theta_{\alpha_i}$.

Nous pouvons donc conclure que l'intervalle ayant la probabilité *a posteriori* maximale ne peut être un de ceux qui ne sont pas choisis entre Θ_{α_i} et $\Theta_{\alpha_{i+1}}$.

Ceci conclut la preuve du théorème général.

Avant de clore cette annexe, nous voudrions mentionner que la méthode de sélection des β_i est faite pour minimiser le nombre de β_i à définir et conserver le résultat. Tant que les conditions de base d'existence des β_i que nous avons utilisés dans la preuve n'est pas modifiée, pour toute modification (généralement afin de la simplifier) de la méthode de sélection des β_i , les résultats du théorème seront maintenus. C'est-à-dire que l'on peut choisir davantage de β_i , en autant qu'il ne puisse pas y avoir de contradiction avec la méthode de sélection présentée ici.

BIBLIOGRAPHIE

- [1] EVANS, M. (1997), Bayesian inference procedures derived via the concept of relative surprise, *Communications in Statistics* **26**, 1125-1143.
- [2] RAMLAU-HANSEN, H. (1988), A solvency study in non-life insurance. Part 1. Analyses of fire, windstorm and glass claims, *Scandinavian Actuarial Journal* **0**, 3-34.
- [3] ROBERT, C. P. (2001), *The Bayesian choice*, 2^e édition, Springer-Verlag, New York.
- [4] VARIAN, H. R. (1975), A Bayesian approach to real estate assessment, *Studies in Bayesian Econometrics and Statistics in Honor of Leonard J. Savage*, 195-208.
- [5] ZELLNER, A. (1986), Bayesian estimation and prediction using asymmetric loss functions, *Journal of the American Statistical Association* **81**, 446-451.