

2411.2646.6

Université de Montréal

**COMPARAISON BAYÉSIENNE DE COÛTS
ENTRE DEUX TRAITEMENTS
POUR DES MÉLANGES DE LOIS**

par

Alain Desgagné

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 1998



CHIL 2641.6

QA
3
U54
1998
V.020

Université de Montréal

COMPARAISON HAYÉSIENNE DE COÛTS
ENTRE DEUX TRAITEMENTS
POUR DES MÉLANGES DE LOIS

Alain Desjardins

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

Atteint par le droit de la propriété intellectuelle
et de la Loi sur l'accès à l'information (S. 30)

Centre de recherche



UNIVERSITÉ DE MONTRÉAL

Université de Montréal

Faculté des études supérieures

Ce mémoire intitulé

**COMPARAISON BAYÉSIENNE DE COÛTS
ENTRE DEUX TRAITEMENTS
POUR DES MÉLANGES DE LOIS**

présenté par

Alain Desgagné

a été évalué par un jury composé des personnes suivantes :

Yves Lepage

(président-rapporteur)

Jean-François Angers

(directeur de recherche)

Jacques LeLorier

(co-directeur)

Louis Doray

(membre du jury)

Mémoire accepté le :

6 octobre 1998

SOMMAIRE

Dans ce mémoire, nous voulons comparer les coûts de deux traitements concurrents. En pharmacoeconomie, il arrive souvent qu'un traitement étudié génère des effets secondaires rares mais coûteux. Il en résulte une distribution des coûts asymétrique, voire même bimodale, avec un mode représentant les effets secondaires. Les techniques statistiques usuelles sont alors inadéquates pour comparer des traitements avec de telles distributions de coûts.

Nous sommes donc intéressés à trouver des méthodes statistiques alternatives. Nous présentons tout d'abord trois méthodes de rééchantillonnage qui permettent d'effectuer adéquatement un test d'hypothèse d'égalité des moyennes. Ce sont des approches fréquentistes non paramétrique pour la première, paramétrique pour la deuxième et un mélange des deux pour la troisième.

L'autre approche est bayésienne. Tout comme la méthode de rééchantillonnage paramétrique, elle requiert la modélisation des données par une distribution paramétrique. Nous présentons le modèle d'un mélange de deux distributions, plus particulièrement des distributions log-normales. Ce type de distribution permet de bien modéliser des données comportant des valeurs extrêmes représentant les coûts des effets secondaires. Nous introduisons différentes méthodes pour estimer les paramètres d'une telle distribution. L'estimateur du maximum de vraisemblance et l'estimateur du maximum de vraisemblance généralisé (valeur qui maximise la densité *a posteriori*) sont les méthodes d'estimation les plus efficaces. L'algorithme EM est utilisé pour trouver ces estimateurs.

L'approche bayésienne consiste à combiner l'information *a priori* sur les paramètres des distributions et l'information fournie par les observations pour déduire une distribution des paramètres *a posteriori*. Ensuite, il est possible d'effectuer un test d'hypothèse d'égalité des moyennes de façon bayésienne en calculant la probabilité *a posteriori* que l'hypothèse H_0 survienne. À partir d'une fonction de coûts adéquate, il est ensuite possible de trancher entre les deux hypothèses H_0 et H_1 . Une autre façon de comparer les coûts des traitements est de calculer la probabilité *a posteriori* qu'un traitement soit moins coûteux que l'autre. Finalement, ces méthodes sont appliquées à des données réelles provenant d'une étude pharmacoéconomique.

REMERCIEMENTS

J'aimerais remercier tout particulièrement mon directeur de recherche, Monsieur Jean-François Angers, qui a été d'une grande disponibilité, d'un commerce toujours agréable et qui sut très bien me guider à travers les méandres d'un mémoire.

J'aimerais également remercier mon codirecteur de recherche, Monsieur Jacques LeLorier, pour son apport pharmacoéconomique au projet, et aussi pour les nombreuses opportunités qu'il m'a permis de saisir.

Je suis aussi reconnaissant au Fonds CRSNG pour la bourse d'études qui m'a été octroyée tout au long de ma maîtrise.

Je suis également reconnaissant à la compagnie pharmaceutique Pfizer, par l'entremise de Madame Louise Perrault, pour son support financier.

Finalement, je ne saurais oublier ma conjointe Anne-Marie qui m'a toujours supporté et même aidé dans la rédaction de ce mémoire. Sa patience et ses encouragements m'ont grandement aidé dans la réalisation de ce projet.

TABLE DES MATIÈRES

| | |
|--|------|
| SOMMAIRE | ii |
| REMERCIEMENTS | iv |
| LISTE DES FIGURES | viii |
| LISTE DES TABLEAUX | ix |
| INTRODUCTION | 1 |
| CHAPITRE 1. APPROCHE CLASSIQUE DES TESTS D’HY- POTHÈSES | 4 |
| 1.1. PRÉSENTATION DU PROBLÈME | 4 |
| 1.2. MÉTHODES STATISTIQUES CLASSIQUES | 8 |
| 1.2.1. Test de Student | 8 |
| 1.2.2. Tests non paramétriques | 10 |
| 1.2.2.1. Test de Mann-Whitney-Wilcoxon | |
| 1.2.2.2. Test de la médiane | |
| 1.2.3. Test de Student sur des données transformées ou tronquées | 13 |
| 1.2.3.1. Transformation des données | |
| 1.2.3.2. Troncation des données | |
| 1.3. MÉTHODE DE RÉÉCHANTILLONNAGE NON PARAMÉTRIQUE | 15 |
| 1.3.1. Algorithme de la méthode | 15 |
| 1.3.2. Nombre de simulations et précision | 19 |

| | |
|---|-----------|
| CHAPITRE 2. ESTIMATION DES PARAMÈTRES DE DISTRIBUTIONS BIMODALES | 23 |
| 2.1. MÉTHODES TRADITIONNELLES D'ESTIMATION DES PARAMÈTRES | 24 |
| 2.1.1. Présentation du modèle de mélange de deux lois | 24 |
| 2.1.2. Méthode graphique | 26 |
| 2.1.3. Méthode des moments | 26 |
| 2.1.4. Méthode du maximum de vraisemblance | 30 |
| 2.2. ALGORITHME EM | 30 |
| 2.2.1. Description de l'algorithme EM | 30 |
| 2.2.2. Application de l'algorithme EM à l'ajustement de distributions bimodales | 32 |
| 2.2.3. Application de l'algorithme EM à des mélanges de lois spécifiques | 35 |
| 2.2.3.1. Mélanges de lois normales | |
| 2.2.3.2. Mélange de lois gamma | |
| 2.2.3.3. Mélange de lois log-normales | |
| 2.2.3.4. Mélange de lois inverse-gamma | |
| 2.2.4. Exemple d'application | 40 |
| 2.3. MÉTHODE DE RÉÉCHANTILLONNAGE PARAMÉTRIQUE | 42 |
| 2.4. MÉTHODE DE RÉÉCHANTILLONNAGE MODIFIÉ | 45 |
| CHAPITRE 3. APPROCHE BAYÉSIENNE | 51 |
| 3.1. TESTS D'HYPOTHÈSES BAYÉSIENS | 52 |
| 3.2. LOIS CONJUGUÉES | 54 |
| 3.3. ESTIMATION DES PARAMÈTRES PAR LA MÉTHODE BAYÉSIENNE | 58 |
| 3.4. TEST D'HYPOTHÈSES UNILATÉRAL AVEC POIDS ET VARIANCES CONNUS | 64 |
| 3.5. TEST D'HYPOTHÈSES BILATÉRAL AVEC POIDS ET VARIANCES CONNUS | 74 |

| | |
|--|------------|
| 3.6. TEST D'HYPOTHÈSES UNILATÉRAL AVEC POIDS ET VARIANCES ALÉATOIRES | 80 |
| 3.7. TEST D'HYPOTHÈSES BILATÉRAL AVEC POIDS ET VARIANCES ALÉATOIRES | 91 |
| 3.8. APPLICATION DES TESTS D'HYPOTHÈSES UNILATÉRAL ET BILATÉRAL SUR LES DONNÉES DE L'ÉTUDE | 97 |
| 3.9. PROBABILITÉ A POSTERIORI QUE LE TRAITEMENT 1 SOIT MOINS COÛTEUX QUE LE TRAITEMENT 2 | 99 |
| CONCLUSION | 104 |
| ANNEXE 1. DÉMONSTRATIONS | 106 |
| Démonstration de l'équation (3.4.14) | 106 |
| Démonstration de l'équation (3.4.22) | 106 |
| Démonstration de l'équation (3.5.2) | 108 |
| Démonstration de l'équation (3.6.12) | 110 |
| Démonstration de l'équation (3.7.1) | 111 |
| ANNEXE 2. PROGRAMMES INFORMATIQUES | 114 |
| RÉFÉRENCES | 138 |

LISTE DES FIGURES

- | | | |
|-----|---|----|
| 1.1 | Histogramme des coûts du traitement à la TMP/SMX | 7 |
| 1.2 | Histogramme des coûts du traitement à la ciprofloxacine | 7 |
| 1.3 | Histogramme des $B=10\ 000$ répliques de la statistique Z (Z^*) sous H_0 calculée avec la méthode de rééchantillonnage non paramétrique | 18 |

LISTE DES TABLEAUX

| | | |
|-----|---|----|
| 1.1 | Coûts (en dollars) des traitements pour chaque patient | 6 |
| 1.2 | Statistiques descriptives des coûts (en dollars) des traitements | 6 |
| 1.3 | Nombre de simulations (B) requis pour atteindre différentes valeurs du coefficient de variation en estimant p par $\hat{p}=0,155$ | 20 |
| 1.4 | Intervalle de confiance de niveau 95 % pour p pour différentes valeurs de B avec $\hat{p}=0,155$ | 20 |
| 1.5 | Résultats des différents tests statistiques discutés dans le chapitre 1 | 21 |
| 2.1 | Ajustement de distributions bimodales pour le traitement à la TMP/SMX | 40 |
| 2.2 | Ajustement de distributions bimodales pour le traitement à la ciprofloxacine | 41 |
| 2.3 | Espérance et variance de distributions log-normale, gamma et inverse-gamma | 42 |
| 2.4 | Résultats des méthodes de rééchantillonnage | 50 |
| 3.1 | Choix des paramètres des lois <i>a priori</i> pour l'ajustement d'un mélange de lois log-normales par la méthode bayésienne | 64 |
| 3.2 | Résultats de l'ajustement d'un mélange de lois log-normales par la méthode bayésienne | 64 |

INTRODUCTION

En pharmacoeconomie, la comparaison des coûts engendrés par la consommation de différents médicaments utilisés concurremment pour le traitement de la même maladie représente un grand intérêt. Le problème est d'autant plus important lorsque les médicaments sont susceptibles d'engendrer des effets secondaires majeurs, ce qui est souvent le cas. La distribution des coûts est alors asymétrique et parfois même bimodale dû à un petit nombre de patients qui génèrent de très grands coûts. Dans ce mémoire, nous utilisons des données provenant d'une étude pharmacoeconomique. Ces données nous serviront d'exemple tout au long de notre étude.

La méthode statistique classique la plus utilisée pour comparer les coûts moyens engendrés par deux médicaments est le test de Student. Si la distribution des coûts de chacun des traitements ne suit pas une loi normale, les tests de rang tels le test de Mann-Whitney-Wilcoxon ou le test de la médiane peuvent constituer une alternative au test de Student. Sinon, nous tentons souvent de normaliser les données par une transformation, le plus souvent logarithmique, ou encore par une troncation des données extrêmes, avant d'appliquer le test de Student.

Toutefois, ces méthodes sont inappropriées lorsque les données sont trop asymétriques. Nous passons en revue ces méthodes dans le premier chapitre et discutons pourquoi elles peuvent être inappropriées. Toujours au chapitre 1, la méthode de rééchantillonnage non paramétrique est présentée comme première méthode alternative.

Le chapitre 2 est consacré à l'ajustement de distributions et à deux versions paramétriques de la méthode de rééchantillonnage. Pour mieux tenir compte de l'asymétrie des données causée par les effets secondaires rares mais coûteux, une loi bimodale est ajustée aux données avec un mode distinct pour ces effets secondaires. Nous voyons les différentes méthodes utilisées pour ajuster un mélange de lois, soient les méthodes graphique, des moments, du maximum de vraisemblance et la méthode EM. Cette dernière méthode, la plus efficace, est ensuite utilisée pour ajuster des mélanges de lois aux données. Le mélange de densités log-normales donnant le meilleur ajustement, il est choisi pour modéliser les données pour chaque médicament.

À partir des lois ajustées, nous procédons à la méthode de rééchantillonnage paramétrique. Cette méthode diffère de la méthode de rééchantillonnage non paramétrique par l'hypothèse faite sur les distributions des données. La méthode de rééchantillonnage non paramétrique suppose que les données proviennent de leur distribution empirique, tandis que la méthode de rééchantillonnage paramétrique suppose que les données proviennent de distributions paramétriques.

Finalement dans le chapitre 2, nous proposons une nouvelle version de la méthode de rééchantillonnage, soit la méthode de rééchantillonnage modifié. C'est une méthode hybride de la méthode non paramétrique et de la méthode paramétrique, où la distribution paramétrique est utilisée pour calculer la probabilité d'observer des coûts plus grands que la plus grande donnée observée.

Enfin une solution bayésienne est décrite dans le chapitre 3. L'approche bayésienne consiste à combiner l'information *a priori* et l'information fournie par les observations pour en déduire une distribution des paramètres *a posteriori*. À partir de cette distribution, nous calculerons la probabilité *a posteriori* que l'hypothèse H_0 survienne et la probabilité *a posteriori* qu'un traitement soit moins coûteux que l'autre.

Mais tout d'abord à la section 3.1, nous présentons les tests d'hypothèses selon l'approche bayésienne. Cela consiste à choisir une fonction de coûts adéquate

qui pénalise les mauvaises décisions. Nous démontrons que la règle de décision qui minimise cette fonction de coûts dépend de la probabilité *a posteriori* que l'hypothèse H_0 survienne.

Nous introduisons à la section 3.2 les familles de lois conjuguées qui nous aideront à choisir les lois *a priori* de façon à ce qu'elles soient de la même famille que les lois *a posteriori*.

À la section 3.3, nous présentons une autre approche bayésienne aux tests d'hypothèses. Il s'agit d'appliquer la méthode de rééchantillonnage paramétrique avec des distributions ajustées aux données à partir de la méthode bayésienne. La méthode bayésienne consiste à appliquer l'algorithme EM tel que vu au chapitre 2, pour trouver les estimateurs qui maximisent la distribution *a posteriori* des paramètres.

Aux sections 3.4 à 3.7, nous présentons les algorithmes pour appliquer les tests d'hypothèses unilatéral et bilatéral selon l'approche bayésienne, selon qu'une partie ou que la totalité des paramètres soient considérés comme des variables aléatoires. Nous appliquons ensuite les algorithmes des sections 3.4 et 3.5 aux données de l'étude et les résultats sont présentés à la section 3.8. Finalement à la section 3.9, nous proposons une méthode bayésienne alternative. Il s'agit de calculer la probabilité *a posteriori* qu'un traitement soit moins coûteux que l'autre.

CHAPITRE 1

APPROCHE CLASSIQUE DES TESTS D'HYPOTHÈSES

Dans ce premier chapitre, nous allons tout d'abord présenter à la section 1.1 l'étude pharmacoéconomique qui a généré le problème statistique que ce mémoire tente de résoudre. Le problème rencontré est le manque de méthodes statistiques appropriées pour comparer les moyennes de coûts de deux traitements avec deux médicaments différents quand les données sont asymétriques. Cette asymétrie est souvent causée par les effets secondaires d'un médicament qui peuvent être rares mais très coûteux.

Nous verrons à la section 1.2 les méthodes statistiques classiques utilisées pour répondre au test d'hypothèses d'égalité des moyennes et pourquoi elles peuvent être inappropriées quand les données sont asymétriques. À la section 1.3, nous présentons la méthode de rééchantillonnage non paramétrique comme première alternative efficace aux méthodes statistiques classiques.

1.1. PRÉSENTATION DU PROBLÈME

Lors d'une étude pharmacoéconomique, 89 patientes souffrant de pyélonéphrite aiguë se voyaient administrer en double aveugle (ni la patiente ni le médecin ne sait quel médicament est administré à la patiente) soit de la TMP/SMX (Triméthoprim/Sulfaméthoxazole) ou de la ciprofloxacine. La pyélonéphrite aiguë est une infection bactérienne du rein et des voies urinaires hautes. Nous voulions étudier les coûts directs associés au traitement de la maladie avec l'un ou l'autre

de ces médicaments. La TMP/SMX est présentement sur le marché à un coût d'acquisition de 8,30\$ pour un traitement de 10 jours, ce qui est considéré peu élevé. Elle présente toutefois l'inconvénient de produire des effets secondaires qui peuvent en quelques rares occasions engendrer des coûts très élevés. Les effets secondaires peuvent être des troubles gastro-intestinaux, des réactions allergiques cutanées ou des complications de la maladie, par exemple des abcès. Un nouveau médicament, la ciprofloxacine, arrive sur le marché à un coût d'acquisition de 77,70\$ pour un traitement de 10 jours, ce qui est presque dix fois plus élevé que la TMP/SMX. Par contre, son taux de succès pour guérir la pyélonéphrite aigüe est plus élevé et les effets secondaires sont moins fréquents et surtout moins coûteux. Le coût d'un traitement inclut le coût d'acquisition du médicament, le coût des visites médicales, les frais de laboratoire, le coût du traitement des effets secondaires (parfois même l'hospitalisation) et le coût d'un traitement alternatif en cas d'échec.

Nous voulons savoir si le coût moyen du traitement à la TMP/SMX est statistiquement différent du coût moyen du traitement à la ciprofloxacine. Nous procédons pour ce faire à un test d'égalité des moyennes avec contre-hypothèse bilatérale, ce qui s'écrit:

$$H_0 : \mathbb{E}(X) = \mathbb{E}(Y) \text{ contre } H_1 : \mathbb{E}(X) \neq \mathbb{E}(Y),$$

où $\mathbb{E}(X)$ et $\mathbb{E}(Y)$ représentent respectivement les espérances des coûts des traitements à la TMP/SMX et à la ciprofloxacine. Généralement en pharmacoeconomie, nous n'utilisons pas la contre-hypothèse unilatérale car nous ne savons pas, avant d'avoir réalisé l'étude, lequel des deux traitements risque d'être le plus cher.

Quarante-quatre patientes sont randomisées dans le groupe de la TMP/SMX et 45 dans le groupe de la ciprofloxacine. Les coûts observés des deux traitements sont présentés en détail au tableau 1.1. Aussi, des statistiques descriptives sont présentées au tableau 1.2. Nous y remarquons que la moyenne et le maximum des coûts du premier traitement (TMP/SMX) sont plus élevés que ceux du deuxième traitement (ciprofloxacine), ce qui s'explique par les effets secondaires coûteux

TABLEAU 1.1. Coûts (en dollars) des traitements pour chaque patient

| | | | | | | | | | | | | | | |
|--------------------------------|-----|-----|-----|-----|-----|-----|-----|------|------|------|------|------|------|-----|
| Traitement à la TMP/SMX | | | | | | | | | | | | | | |
| 127 | 127 | 127 | 128 | 148 | 149 | 162 | 191 | 191 | 192 | 194 | 200 | 211 | 216 | 216 |
| 219 | 219 | 227 | 227 | 232 | 232 | 242 | 257 | 261 | 286 | 298 | 308 | 309 | 323 | 329 |
| 340 | 341 | 379 | 380 | 392 | 484 | 513 | 994 | 1328 | 1867 | 1891 | 1921 | 2319 | 6182 | |
| Traitement à la ciprofloxacine | | | | | | | | | | | | | | |
| 192 | 192 | 192 | 192 | 200 | 200 | 212 | 213 | 225 | 257 | 257 | 257 | 258 | 259 | 262 |
| 264 | 266 | 266 | 266 | 267 | 267 | 270 | 276 | 279 | 279 | 288 | 289 | 289 | 290 | 291 |
| 291 | 291 | 298 | 304 | 308 | 311 | 312 | 314 | 325 | 377 | 381 | 438 | 461 | 465 | 743 |

TABLEAU 1.2. Statistiques descriptives des coûts (en dollars) des traitements

| Traitement | Nombre de patientes | Moyenne | Médiane | Écart type | Min. | Max. |
|----------------|---------------------|---------|---------|------------|------|------|
| TMP/SMX | 44 | 588 | 250 | 1018 | 127 | 6182 |
| ciprofloxacine | 45 | 292 | 276 | 94 | 192 | 743 |

de la TMP/SMX. Nous remarquons toutefois que la médiane et le minimum des coûts du premier traitement sont inférieurs à ceux du deuxième traitement, ce qui s'explique par le coût d'acquisition de la TMP/SMX presque dix fois moins élevé que celui de la ciprofloxacine. De plus, l'écart type des coûts du premier traitement est près de 11 fois plus élevé que celui du deuxième traitement, ce qui s'explique aussi par les effets secondaires coûteux reliés à la TMP/SMX. Les histogrammes des coûts des premier et deuxième traitements sont présentés respectivement aux figures 1.1 et 1.2. Nous pouvons y voir l'asymétrie des coûts des traitements causée par les effets secondaires, de façon plus marquée pour le traitement à la TMP/SMX.

À la lumière de l'analyse des données et statistiques, la conclusion du test d'égalité des moyennes dépendra de la réponse à cette question: Est-ce que le coût d'acquisition de la ciprofloxacine presque dix fois plus élevé que celui de la TMP/SMX est compensé par son plus haut taux de succès et par l'incidence et la sévérité de ses effets secondaires moins élevées? Dans la prochaine section, les méthodes statistiques classiques utilisées pour le test sont discutées.

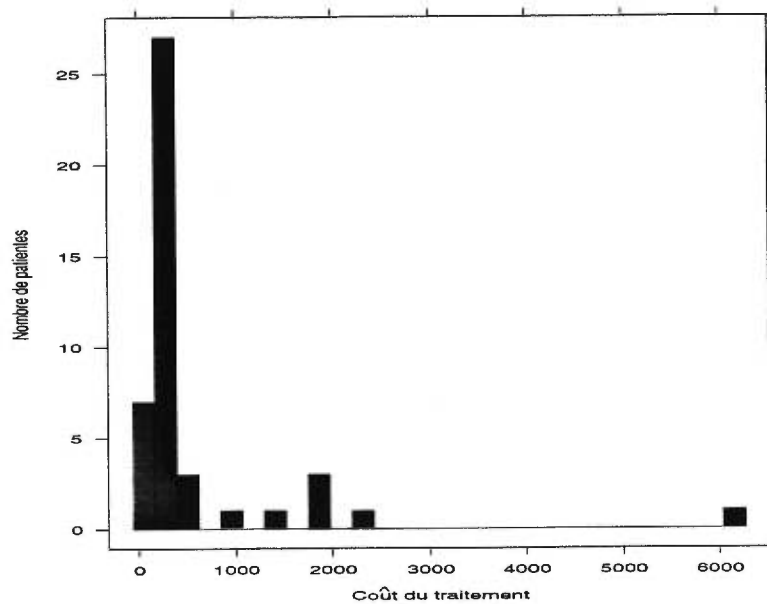


FIGURE 1.1. Histogramme des coûts du traitement à la TMP/SMX

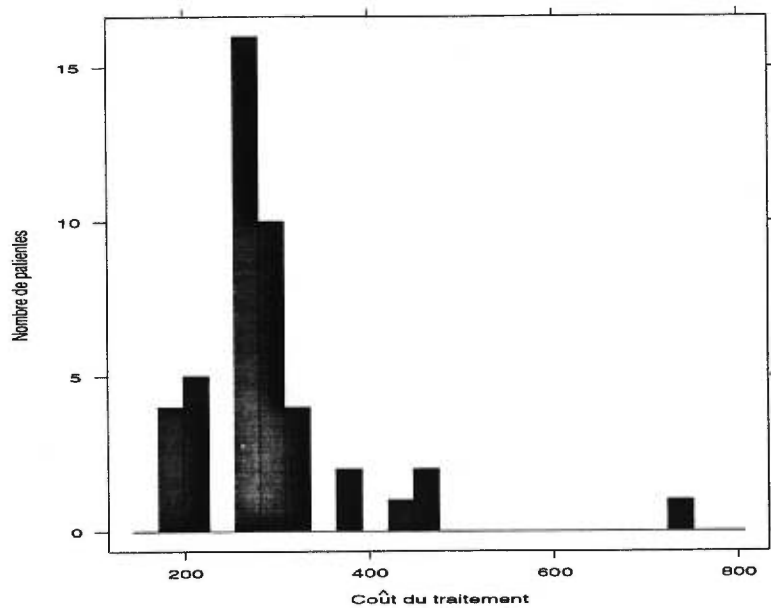


FIGURE 1.2. Histogramme des coûts du traitement à la ciprofloxacine

1.2. MÉTHODES STATISTIQUES CLASSIQUES

1.2.1. Test de Student

La méthode la plus populaire pour choisir entre H_0 et H_1 est le test de Student. Revoyons les étapes de ce test puisque les méthodes de rééchantillonnage non paramétrique et paramétrique présentées plus tard sont construites de la même façon. Nous divisons le test de Student en quatre étapes.

Définissons d'abord X et Y comme étant les variables aléatoires indépendantes représentant respectivement les coûts par patient pour les traitements à la TMP/-SMX et à la ciprofloxacine. L'échantillon des coûts observés est représenté par (x_1, \dots, x_m) pour le premier traitement et par (y_1, \dots, y_n) pour le deuxième traitement, où m et n représentent le nombre d'observations de chaque échantillon. Dans l'étude considérée, $m = 44$ et $n = 45$. Nous supposons que les $m + n$ observations sont indépendantes les unes des autres.

La première étape du test de Student consiste à faire un présupposé sur les distributions inconnues des variables aléatoires X et Y . Nous supposons que X suit une loi normale de moyenne μ_x et de variance σ_x^2 et que Y suit une loi normale de moyenne μ_y et de variance σ_y^2 , où μ_x , μ_y , σ_x^2 et σ_y^2 sont inconnues. Nous notons ces distributions comme suit:

$$X \sim N(\mu_x, \sigma_x^2) \text{ et } Y \sim N(\mu_y, \sigma_y^2).$$

Les histogrammes des coûts des figures 1.1 et 1.2 suggèrent que les distributions des coûts des deux traitements ne suivent pas des lois normales, ce qui est confirmé par un test de Kolmogorov-Smirnov, décrit par Hogg et Tanis (1988, section 9.7), qui donne une valeur- p de moins de 0,001 pour le premier traitement et de 0,005 pour le deuxième traitement. L'hypothèse de normalité n'est donc pas satisfaite ni pour X ni pour Y .

La deuxième étape consiste à choisir une statistique nous permettant de répondre au test $H_0 : \mu_x = \mu_y$ contre $H_1 : \mu_x \neq \mu_y$. La statistique que nous

choisissons est notée et définie par

$$Z = \frac{(\bar{x} - \bar{y})}{\sqrt{(s_x^2/m) + (s_y^2/n)}},$$

où

$$\bar{x} = \frac{1}{m} \sum_{i=1}^m x_i, \quad s_x^2 = \frac{1}{m-1} \sum_{i=1}^m (x_i - \bar{x})^2,$$

$$\bar{y} = \frac{1}{n} \sum_{i=1}^n y_i \quad \text{et} \quad s_y^2 = \frac{1}{n-1} \sum_{i=1}^n (y_i - \bar{y})^2.$$

La statistique Z est une variable aléatoire et une valeur possible est z_{obs} , soit Z calculée à partir des deux échantillons observés. Dans l'étude, $z_{obs}=1,92$.

La statistique du test de Student serait différente si nous faisons l'hypothèse que les variances des deux distributions sont égales. Nous ne ferons toutefois pas cette hypothèse pour ne pas perdre de généralité. De plus, en pharmacoéconomie, nous rencontrons souvent le type d'étude où un médicament génère des effets secondaires plus coûteux que l'autre, ce qui produit des variances différentes entre les deux traitements. Dans l'étude, il est visible graphiquement que les variances sont statistiquement différentes, ce qui est confirmé par un test de Fisher, décrit dans Hogg et Tanis (1988, section 7.4), qui donne une valeur- p de moins de 0,001.

La troisième étape consiste à trouver la distribution de la statistique Z sous l'hypothèse $H_0 : \mu_x = \mu_y$. Comme nous avons supposé la normalité des distributions, nous pouvons montrer que la statistique Z sous H_0 suit approximativement une loi de Student. Le nombre de degrés de liberté peut être approximé par la formule suivante:

$$d.l. \doteq \frac{(s_x^2/m + s_y^2/n)^2}{\frac{1}{m-1} (s_x^2/m)^2 + \frac{1}{n-1} (s_y^2/n)^2}.$$

Dans l'étude, le nombre de degrés de liberté est 43,7.

Finalement, la quatrième étape consiste à calculer la valeur- p . De façon générale,

$$p = \mathbb{P}_{H_0} [|Z| \geq |z_{obs}|]$$

où $|Z|$ est la valeur absolue de Z . Plus particulièrement pour le test de Student,

$$p = 2 \times \mathbb{P}[t_{d.l.} \geq |z_{obs}|],$$

et pour l'étude, nous avons:

$$p = 2 \times \mathbb{P}[t_{43,7} \geq 1,92] = 0,061.$$

Avec un niveau de tolérance de 5% pour l'erreur de type I, nous ne pourrions pas conclure à une différence entre les coûts des deux traitements, mais nous le pourrions avec un niveau de tolérance de 10%. La valeur- p du test de Student avec contre-hypothèse unilatérale est tout simplement la moitié du p du test bilatéral, puisque la loi de Student est symétrique.

Le test de Student est approprié si les données sont distribuées selon une loi normale. De plus, il permet d'ajuster pour des variances inégales. Toutefois, si la taille des échantillons est assez grande, le test de Student sera robuste à la non-normalité des données. En effet, par le théorème limite central tel que qu'énoncé dans Hogg et Tanis (1988, section 5.3), \bar{X} tend vers une $N\left(\mathbb{E}(X), \frac{1}{m}\text{Var}(X)\right)$ pour toute distribution de X , si les X_i sont indépendants et identiquement distribués et si la variance de X existe. Toutefois, pour des tailles d'échantillon plus petites, le test de Student est moins robuste à la non-normalité des données, particulièrement lorsqu'elles sont très asymétriques. Il n'est donc pas recommandé d'utiliser le test de Student pour notre étude (voir Bickel et Doksum (1977, section 6.4)).

1.2.2. Tests non paramétriques

Les tests de rang tels le test de Mann-Whitney-Wilcoxon et le test de la médiane, décrits par Randles et Wolfe (1979, sections 3.4 et 8.1), sont souvent des méthodes alternatives utilisées lorsque le test de Student n'est pas approprié. Ce sont deux méthodes dites non paramétriques ou libres de distribution. Ces méthodes nécessitent toutefois l'hypothèse que si les distributions des deux populations diffèrent, c'est seulement par leur paramètre de position. Nous devons

donc supposer que les formes et variances des distributions sont égales, ce qui n'est toutefois pas le cas dans l'étude, où $s_x^2 = 1018^2$ et $s_y^2 = 94^2$. Rappelons que le test de Fisher donne une valeur- p de moins de 0,001. Les hypothèses testées sont en fait:

$$H_0 : M_x = M_y \text{ contre } H_1 : M_x \neq M_y$$

où M_x et M_y sont les médianes respectives des premier et deuxième traitements. Ces méthodes de rang sont appropriées pour tester la différence entre deux médianes plutôt qu'entre deux moyennes, ce qui n'est pas nécessairement la même chose. Par exemple, dans le cas de distributions très asymétriques, la médiane diffère significativement de la moyenne. Ces méthodes réduisent considérablement l'impact des données extrêmes en leur assignant un rang. Toutefois, en pharmacoeconomie, ces données extrêmes causées par les effets secondaires coûteux sont de très grande importance. Les méthodes de rang ne sont donc pas appropriées pour notre étude. Voyons quand même ces deux méthodes.

1.2.2.1. *Test de Mann-Whitney-Wilcoxon*

Le test de Mann-Whitney-Wilcoxon consiste à combiner les deux échantillons (x_1, \dots, x_m) et (y_1, \dots, y_n) en un seul puis à ordonner de la plus petite à la plus grande les $m + n$ observations. Nous leur assignons un rang de 1 à $m + n$. S'il y a plusieurs observations ayant la même valeur, nous leur assignons la moyenne des rangs qu'elles auraient occupés si elles n'avaient pas eu la même valeur. Nous calculons S comme étant la somme des rangs des observations de la première population, soit la somme des rangs des x_i , $i = 1, \dots, m$. Si m et n sont toutes deux assez grandes (> 20 , voir Daniel (1990, section 3.1)), nous pouvons appliquer le théorème limite central pour approximer la distribution de S sous H_0 . Si c'est le cas, alors

$$Z = \frac{S - m(m + n + 1)/2}{\sqrt{m n(m + n + 1)/12}}$$

suit approximativement une distribution normale de moyenne 0 et de variance 1 lorsque H_0 est vraie. Nous pouvons corriger par un facteur d'ajustement si des

égalités de rangs se produisent, mais l'effet sur la valeur- p est négligeable, à moins d'avoir une large proportion d'égalité.

Pour l'étude, $S=1914,5$, $m=44$, $n=45$ et $z_{obs}=-0,54$. Nous calculons la valeur- p par la formule suivante:

$$p = 2 \times \mathbb{P}[N(0, 1) \geq |z_{obs}|] = 0,589.$$

1.2.2.2. Test de la médiane

Le test de la médiane permet d'effectuer le même test que celui de Mann-Whitney-Wilcoxon, soit $H_0 : M_x = M_y$ contre $H_1 : M_x \neq M_y$. Il consiste à combiner les deux échantillons en un seul puis à compter le nombre, disons V , de valeurs de X plus petites que la médiane des échantillons combinés. Rappelons que la médiane est la $\frac{m+n+1}{2}$ ième plus petite valeur si $m+n$ est impair ou la moyenne entre les $\frac{m+n}{2}$ et $\frac{m+n}{2} + 1$ ième plus petites valeurs si $m+n$ est pair.

Nous pouvons encore une fois approximer, par le théorème limite central, la distribution de V sous H_0 , si m et n sont assez grands. Si c'est le cas, alors

$$Z = \begin{cases} \frac{V - \frac{(m+n-1)m}{(m+n)2}}{\sqrt{\frac{(m+n+1)mn}{4(m+n)^2}}} & ; \text{ si } m+n \text{ est impair,} \\ \frac{V - \frac{m}{2}}{\sqrt{\frac{mn}{4(m+n-1)}}} & ; \text{ si } m+n \text{ est pair,} \end{cases}$$

suit approximativement une distribution normale de moyenne 0 et de variance 1 lorsque H_0 est vraie.

Pour l'étude, la médiane des échantillons combinés est 267, $V = 24$ et $z_{obs}=0,95$. Nous calculons la valeur- p par la formule suivante:

$$p = 2 \times \mathbb{P}[N(0, 1) \geq |z_{obs}|] = 0,342.$$

Les valeurs- p de ces deux tests de rang sont assez élevées et ne nous permettent pas de conclure à une différence entre les médianes des coûts des deux traitements.

1.2.3. Test de Student sur des données transformées ou tronquées

1.2.3.1. Transformation des données

Lorsque le test de Student n'est pas approprié à cause de la non-normalité des données, une alternative souvent utilisée est de transformer les données dans le but de rendre les distributions normales. Cette méthode est décrite par Zhou, Melfi et Hui (1997). La transformation de Box-Cox telle que décrite en détail par Box et Cox (1964) est définie par

$$z_i = \begin{cases} \frac{x_i^\alpha - 1}{\alpha} & ; \text{ si } \alpha \neq 0, \\ \log x_i & ; \text{ si } \alpha = 0, \end{cases}$$

et elle est appliquée en choisissant α de façon à rendre la distribution des x_i le plus près possible d'une loi normale. La transformation logarithmique, lorsque $\alpha = 0$, est souvent utilisée pour essayer de normaliser des données asymétriques. Dans ce cas, la première hypothèse à satisfaire est la log-normalité des données originales des deux échantillons, soit

$$X_i \sim LN(\mu_x, \sigma_x^2) \Leftrightarrow \log X_i \sim N(\mu_x, \sigma_x^2), i = 1, \dots, m$$

et

$$Y_i \sim LN(\mu_y, \sigma_y^2) \Leftrightarrow \log Y_i \sim N(\mu_y, \sigma_y^2), i = 1, \dots, n.$$

De plus, comme

$$\mathbb{E}(X) = \exp(\mu_x + \sigma_x^2/2) \text{ et } \mathbb{E}(Y) = \exp(\mu_y + \sigma_y^2/2),$$

nous devons tester

$$H_0 : \mathbb{E}(X) = \mathbb{E}(Y), \text{ soit } H_0 : \exp(\mu_x + \sigma_x^2/2) = \exp(\mu_y + \sigma_y^2/2).$$

En appliquant le test de Student sur les données log-transformées, nous testons plutôt

$$H_0 : \mathbb{E}(\log X) = \mathbb{E}(\log Y), \text{ soit } H_0 : \mu_x = \mu_y,$$

ce qui est équivalent au bon test uniquement si $\sigma_x^2 = \sigma_y^2$. La deuxième hypothèse à satisfaire lorsque nous voulons transformer les données est donc de s'assurer que $\sigma_x^2 = \sigma_y^2$, soit $\text{Var}(\log X) = \text{Var}(\log Y)$, c'est-à-dire s'assurer que les variances des

données log-transformées soient égales, sans quoi nous ne testons pas la bonne hypothèse. Un test de Kolmogorov-Smirnov pour tester si la distribution log-transformée est normale donne une valeur- p de 0,022 pour le premier traitement et de 0,097 pour le deuxième traitement. De plus, les variances des données log-transformées sont très différentes, soient 0,782 et 0,068 pour les premier et deuxième traitements respectivement, ce qui est confirmé par un test de Fisher qui donne une valeur- p de moins de 0,001. Les hypothèses requises pour utiliser le test de Student sur les données log-transformées ne sont donc pas remplies. Si nous appliquons quand même le test de Student aux données log-transformées, nous trouvons une statistique de 1,22 et une valeur- p de 0,229.

Les deux mêmes hypothèses doivent être remplies pour utiliser le test de Student sur les données transformées par Box-Cox en général, c'est-à-dire

$$(1) \quad \begin{aligned} \frac{X_i^\alpha - 1}{\alpha} &\sim N(\mu_x, \sigma_x^2), i = 1, \dots, m, \\ \frac{Y_i^\alpha - 1}{\alpha} &\sim N(\mu_y, \sigma_y^2), i = 1, \dots, n \end{aligned}$$

$$(2) \quad \sigma_x^2 = \sigma_y^2 \Leftrightarrow \text{Var}(X^\alpha) = \text{Var}(Y^\alpha).$$

Pour les transformations de Box-Cox en général, aucun α ne permet de trouver une transformation qui normalise les données de l'étude tout en produisant des variances égales des données transformées.

1.2.3.2. Troncation des données

Une alternative au test de Student proposée par Coyle (1996) est de tronquer les données extrêmes dans le but de rendre les distributions symétriques. Cette méthode est acceptable si ces données extrêmes sont dues à des erreurs ou encore si nous ne voulons pas les considérer dans l'étude pour une raison quelconque. Toutefois, enlever ces données n'est pas souhaitable dans le contexte d'une étude

pharmacoéconomique où les coûts élevés des effets secondaires jouent un rôle très important.

Afin de comparer le résultat de cette méthode à ceux des autres tests, nous appliquons quand même le test de Student sur les données tronquées en supposant les données extrêmes peu importantes. Toutefois quelles données devons-nous tronquer? Enlever une observation ou la garder peut faire varier considérablement la valeur- p . Nous avons enlevé des valeurs extrêmes de chaque échantillon jusqu'à ce que la valeur- p du test de normalité de Kolmogorov-Smirnov soit supérieur à 0,05. Nous avons ainsi enlevé six observations pour le premier traitement (de 1328\$ à 6182\$) et une observation pour le deuxième traitement (743\$). Le test de Kolmogorov-Smirnov donne respectivement des valeurs- p de 0,203 et de 0,136 pour les premier et deuxième traitements, ce qui indique des distributions normales. Nous appliquons le test de Student sur ces données tronquées et nous trouvons une statistique de -0,33 et une valeur- p de 0,745. Si nous enlevons une autre observation pour le premier traitement (994\$) et que nous appliquons le test de Student, nous trouvons une statistique de -1,53 et une valeur- p de 0,132, ce qui est très loin du $p=0,745$ trouvé avec une donnée de plus.

Le test de Student n'est donc pas recommandé, ni sur les données originales, ni sur les données transformées ou tronquées, quand les données sont asymétriques et la taille de l'échantillon de petite à moyenne. Aucune méthode statistique classique ne répond adéquatement au test d'égalité des moyennes lorsque les données sont très asymétriques. Dans la prochaine section, nous voyons une méthode alternative adéquate et relativement facile à appliquer.

1.3. MÉTHODE DE RÉÉCHANTILLONNAGE NON PARAMÉTRIQUE

1.3.1. Algorithme de la méthode

La méthode de rééchantillonnage non paramétrique pour le test bilatéral d'égalité des moyennes:

$$H_0 : \mathbb{E}(X) = \mathbb{E}(Y) \text{ contre } H_1 : \mathbb{E}(X) \neq \mathbb{E}(Y),$$

est construite de la même façon que le test de Student présenté à la section 1.2.1. Cette méthode est bien décrite par Efron et Tibshirani (1993, section 16.2). Pour les mêmes raisons que le test de Student, nous ne faisons pas l'hypothèse que les variances sont égales. La méthode est décrite en quatre étapes.

La première étape consiste à faire une hypothèse sur les distributions inconnues des variables aléatoires X et Y . Il est démontré par Hall (1992, section 1.2) que la distribution empirique converge asymptotiquement vers la vraie distribution. Alors nous supposons que les observations proviennent de leur distribution empirique. Cette méthode est assurément précise lorsque la taille des échantillons tend vers l'infini. Les fonctions de densité supposées de X et Y sont donc définies par:

$$\mathbb{P}(x) = 1/m \text{ pour } x \in \{x_1, \dots, x_m\}$$

et

$$\mathbb{P}(y) = 1/n \text{ pour } y \in \{y_1, \dots, y_n\}.$$

La deuxième étape consiste à choisir une statistique nous permettant de répondre au test $H_0 : \mathbb{E}(X) = \mathbb{E}(Y)$ contre $H_1 : \mathbb{E}(X) \neq \mathbb{E}(Y)$. Plusieurs statistiques peuvent être choisies, mais comme Efron et Tibshirani (1993, section 16.2) le mentionnent, une plus grande précision peut être obtenue en choisissant la statistique du test de Student, soit

$$Z = \frac{(\bar{x} - \bar{y})}{\sqrt{(s_x^2/m) + (s_y^2/n)}}.$$

La statistique Z est une variable aléatoire et une valeur possible est z_{obs} . Dans l'étude, $z_{obs}=1,92$.

La troisième étape consiste à approximer par B simulations la distribution de la statistique Z sous l'hypothèse $H_0 : \mu_x = \mu_y$. Chaque simulation b ($b = 1, \dots, B$) consiste à générer aléatoirement deux échantillons, notés $(x_1^{*b}, \dots, x_m^{*b})$

et $(y_1^{*b}, \dots, y_n^{*b})$, à partir des distributions empiriques telles que mentionnées dans la première étape, et ce sous l'hypothèse H_0 .

Pour y parvenir, nous tirons au hasard m observations, avec remise, à partir de l'échantillon original (x_1, \dots, x_m) afin de générer $(x_1^{*b}, \dots, x_m^{*b})$. Ensuite pour le deuxième traitement, nous tirons au hasard n observations, avec remise, à partir de $(y_1 + \bar{x} - \bar{y}, \dots, y_n + \bar{x} - \bar{y})$, soit l'échantillon original translaté de $(\bar{x} - \bar{y})$ afin de satisfaire l'hypothèse H_0 . En effet, $\bar{x} - \bar{y} = \mathbb{E}(X) - \mathbb{E}(Y)$ selon l'hypothèse faite à la première étape. Nous obtenons ainsi $(y_1^{*b}, \dots, y_n^{*b})$. Notons que pour satisfaire H_0 , nous aurions pu translater (x_1, \dots, x_m) de $(\bar{y} - \bar{x})$ au lieu de translater (y_1, \dots, y_n) , sans changer la statistique Z et la suite du test. La réplique de la statistique Z , soit

$$Z^{*b} = \frac{(\bar{x}^{*b} - \bar{y}^{*b})}{\sqrt{(s_x^{2*b}/m) + (s_y^{2*b}/n)}},$$

est alors calculée à partir des deux nouveaux échantillons, où

$$\begin{aligned} \bar{x}^{*b} &= \frac{1}{m} \sum_{i=1}^m x_i^{*b}, \quad s_x^{2*b} = \frac{1}{m-1} \sum_{i=1}^m (x_i^{*b} - \bar{x}^{*b})^2, \\ \bar{y}^{*b} &= \frac{1}{n} \sum_{i=1}^n y_i^{*b} \quad \text{et} \quad s_y^{2*b} = \frac{1}{n-1} \sum_{i=1}^n (y_i^{*b} - \bar{y}^{*b})^2. \end{aligned}$$

Cette simulation est faite indépendamment pour $b = 1, \dots, B$ afin d'obtenir Z^{*1}, \dots, Z^{*B} . Un histogramme avec $B = 10\,000$ répliques de la statistique Z sous H_0 (avec les observations de l'étude considérée) apparaît à la figure 1.3, ce qui permet d'illustrer la forme de la densité de Z sous l'hypothèse $H_0 : \mathbb{E}(X) = \mathbb{E}(Y)$. La densité de Z est asymétrique car celle des coûts du premier traitement est très asymétrique.

La quatrième étape consiste à calculer la valeur- p par la formule suivante:

$$p = \mathbb{P}_{H_0} [|Z| \geq |z_{obs}|].$$

Puisque la distribution de Z sous H_0 est approximée par simulation, la valeur- p sera aussi approximée par

$$\hat{p} = \# \{|Z^*| \geq |z_{obs}|\} / B,$$

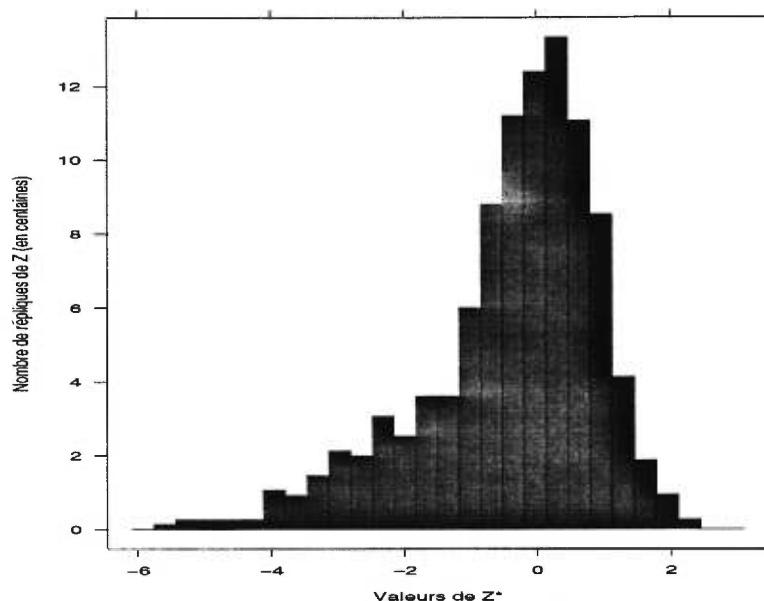


FIGURE 1.3. Histogramme des $B=10\ 000$ répliques de la statistique Z (Z^*) sous H_0 calculée avec la méthode de rééchantillonnage non paramétrique

où $\#\{S\}$ est la cardinalité de l'ensemble S . Dans l'étude, pour $B = 10\ 000$ simulations, 1 516 valeurs de Z^* sont plus petites que $-|z_{obs}| = -1,92$ tandis que 34 valeurs de Z^* sont plus grandes que $|z_{obs}| = 1,92$. Alors

$$\hat{p} = (1\ 516 + 34)/10\ 000 = 0,155.$$

Si nous voulons effectuer le test unilatéral $H_0 : \mathbb{E}(X) = \mathbb{E}(Y)$ contre $H_1 : \mathbb{E}(X) > \mathbb{E}(Y)$, alors

$$\begin{aligned} \hat{p} &= \begin{cases} \#\{Z^* \geq z_{obs}\}/B & \text{si } z_{obs} > 0, \\ \#\{Z^* \leq z_{obs}\}/B & \text{si } z_{obs} < 0 \end{cases} \\ &= 0,0034, \end{aligned}$$

ce qui est très concluant contrairement au test bilatéral. Nous pouvons trouver l'algorithme de la méthode de rééchantillonnage non paramétrique, ainsi que celui des méthodes présentées subséquemment dans ce mémoire, en annexe 2 sous forme de programme informatique. Le langage utilisé est Xlipstat et est bien décrit par Tierney (1990).

En pharmacoéconomie, le test bilatéral est habituellement utilisé car les deux médicaments étudiés ont le potentiel d’avoir un coût moyen de traitement plus élevé que l’autre médicament, même pour un placebo. Dans le cas du test de Student, la valeur- p du test bilatéral est le double de la valeur- p du test unilatéral, puisque la distribution de la statistique Z supposée par ce test, soit la loi de Student, est symétrique. Cependant, la valeur- p du test bilatéral peut être beaucoup plus que le double de celle du test unilatéral lorsque la distribution de la statistique supposée par le test tient compte de l’asymétrie des données, ce qui est le cas avec la méthode de rééchantillonnage. Dans l’étude, la valeur- p obtenue avec la méthode de rééchantillonnage passe de 0,0034 pour le test unilatéral à 0,155 pour le test bilatéral. Le choix de la contre-hypothèse est d’autant plus important lorsque les données sont asymétriques. Dans l’étude considérée, le test unilatéral n’est pas approprié. Desgagné, Castilloux, Angers et LeLorier (1998) discutent de cet aspect de la pharmacoéconomie.

La méthode de rééchantillonnage non paramétrique est une alternative au test de Student et elle est recommandée lorsque les données sont asymétriques et que la taille des échantillons est modérée. C’est une méthode qui a l’avantage de tenir compte de l’asymétrie des données, sans faire d’hypothèse sur les distributions ou l’égalité des variances. De plus, la méthode de rééchantillonnage tient compte de la vraie valeur des observations originales.

1.3.2. Nombre de simulations et précision

Dans l’étude, nous avons effectué $B = 10\,000$ simulations. Toutefois, nous pouvons nous demander combien de simulations devons-nous effectuer pour avoir une précision satisfaisante? Efron et Tibshirani (1993, section 15.2) démontrent que $B \cdot \hat{p}$ suit une distribution binômiale, notée par $Bin(B, p)$. Il en découle ceci:

$$\mathbb{E}(\hat{p}) = p \text{ et } \text{Var}(\hat{p}) = p(1 - p)/B,$$

TABLEAU 1.3. Nombre de simulations (B) requis pour atteindre différentes valeurs du coefficient de variation en estimant p par $\hat{p}=0,155$

| | | | | |
|-----------------|------|-------|-------|--------|
| $CV_B(\hat{p})$ | 0,10 | 0,05 | 0,025 | 0,01 |
| B | 545 | 2 181 | 8 723 | 54 516 |

TABLEAU 1.4. Intervalle de confiance de niveau 95 % pour p pour différentes valeurs de B avec $\hat{p}=0,155$

| | | | | |
|-----------|---------------|---------------|---------------|---------------|
| B | 1 000 | 5 000 | 10 000 | 100 000 |
| I.C. 95 % | (0,133;0,177) | (0,145;0,165) | (0,148;0,162) | (0,153;0,157) |

où p est la vraie valeur- p . Une mesure de la précision de l'estimation de p par \hat{p} est alors donnée par le coefficient de variation de \hat{p} , soit

$$CV_B(\hat{p}) = \frac{\sqrt{\text{Var}(\hat{p})}}{\mathbb{E}(\hat{p})} = \left[\frac{(1-p)}{B \cdot p} \right]^{1/2}.$$

Quand B ou p augmente, le coefficient de variation de \hat{p} diminue et la précision augmente. $CV_B(\hat{p})$ est estimé en approximant p par \hat{p} . Le tableau 1.3 donne le nombre de simulations requis pour atteindre différentes valeurs du coefficient de variation en estimant p par $\hat{p}=0,155$.

Une autre façon de mesurer la précision est de construire un intervalle de confiance pour la valeur- p en utilisant l'approximation suivante:

$$\hat{p} \approx N(p, p(1-p)/B).$$

Un intervalle de confiance avec un niveau de confiance de 95 % pour p est alors donné par:

$$\hat{p} \pm 1,96\sqrt{\hat{p}(1-\hat{p})/B}.$$

Le tableau 1.4 donne l'intervalle de confiance de niveau 95 % pour p pour différentes valeurs de B avec $\hat{p}=0,155$. Nous choisirons donc le nombre de simulations selon la précision souhaitée.

Dans ce chapitre, nous avons présenté et analysé les données de l'étude pharmacoéconomique. Nous avons vu que le premier traitement a un coût d'acquisition considéré peu élevé, mais un taux de succès moyen et des effets secondaires coûteux. Le deuxième traitement a un coût d'acquisition presque dix fois plus

TABLEAU 1.5. Résultats des différents tests statistiques discutés dans le chapitre 1

| | Test de Student | M-W-Wilcoxon | Test de la médiane | Log-transf. | Troncation de données | Rééchant. non para. |
|-------------|-----------------|--------------|--------------------|-------------|-----------------------|---------------------|
| stat. | 1,92 | -0,54 | 0,95 | 1,22 | -0,33 | 1,92 |
| valeur- p | 0,061 | 0,589 | 0,342 | 0,229 | 0,745 | 0,155 |

élevé que le premier, mais un taux de succès relativement élevé et des effets secondaires moins fréquents et surtout moins coûteux que le premier.

Les méthodes statistiques classiques ainsi que la méthode de rééchantillonnage non paramétrique ont été présentées et appliquées sur les données de l'étude. Le tableau 1.5 rappelle les résultats pour le test bilatéral. Nous y voyons des résultats fort variés. La vraie réponse se situe entre les résultats du test de Student sur les données originales ($p=0,061$) et sur les données log-transformées ($p=0,229$). L'un surestime l'impact des valeurs extrêmes, l'autre les sous-estime. Le résultat de la méthode de rééchantillonnage non paramétrique ($p=0,155$) reflète bien sa capacité à mieux estimer l'impact des valeurs extrêmes.

Le test de Student requiert l'hypothèse de normalité des données. Le test de Mann-Whitney-Wilcoxon et le test de la médiane supposent que les distributions ont la même forme et la même variance. De plus, elles considèrent le rang des données au lieu de leur vraie valeur, ce qui enlève du poids aux données extrêmes. Le test de Student sur les données transformées en utilisant les transformations de Box-Cox suppose que les données transformées ont une distribution normale avec des variances égales, sans quoi nous ne testons pas la bonne hypothèse. Le test de Student sur les données tronquées suppose que ces données extrêmes sont dues à des erreurs ou qu'elles ne sont pas importantes. Le résultat du tableau 1.5 a été trouvé en tronquant les six plus grandes observations du premier traitement et la plus grande du deuxième traitement.

Finalement, la méthode de rééchantillonnage non paramétrique est une méthode appropriée pour comparer deux moyennes de coût de traitement de deux médicaments, particulièrement quand les données sont asymétriques. Aucune hypothèse de distribution des données ou d'égalité des variances n'est requise et le

test est basé sur les données originales. De plus, la méthode de rééchantillonnage non paramétrique tient compte de l'asymétrie des données. Cependant, la taille des échantillons ne doit pas être trop petite, sinon une méthode de rééchantillonnage paramétrique sera plus appropriée.

CHAPITRE 2

ESTIMATION DES PARAMÈTRES DE DISTRIBUTIONS BIMODALES

Afin d'utiliser des méthodes paramétriques pour les tests d'hypothèses, telles les méthodes de rééchantillonnage paramétrique et de rééchantillonnage modifié présentées à la fin du chapitre 2 et l'approche bayésienne présentée au chapitre 3, nous nous intéressons dans ce chapitre à l'ajustement de lois de probabilité.

Nous avons vu que les histogrammes des coûts des deux traitements, présentés aux figures 1.1 et 1.2, suggéraient une loi bimodale pour le premier traitement et peut-être même pour le deuxième traitement, le deuxième mode résultant des effets secondaires. L'ajustement de lois unimodales à ces données risque d'être inadéquat, même si nous choisissons des lois dont l'asymétrie vers la droite est marquée, telles les lois log-normale et gamma. Nous ajusterons plutôt des lois bimodales, qui pourront se réduire à des lois unimodales si cela est plus approprié.

À la section 2.1, nous présentons le modèle de mélange de deux lois et nous y calculons l'espérance et la variance. Ensuite, nous voyons les différentes méthodes traditionnelles d'estimation des paramètres, soient la méthode graphique, la méthode des moments et la méthode du maximum de vraisemblance.

À la section 2.2, nous voyons l'algorithme EM que nous appliquons à l'ajustement de distributions bimodales et plus particulièrement aux mélanges de lois normales, gamma, log-normales et inverse-gamma. Nous ajustons ensuite ces modèles aux données de l'étude.

Finalement, aux sections 2.3 et 2.4, nous utilisons le modèle de mélange de deux lois ajusté aux données de l'étude pour présenter deux versions paramétriques de la méthode de rééchantillonnage pour les tests d'hypothèses, soient les méthodes de rééchantillonnage paramétrique et de rééchantillonnage modifié. La première méthode diffère du rééchantillonnage non paramétrique du fait que nous utilisons une distribution paramétrique des données plutôt qu'empirique. La deuxième méthode est un hybride de la méthode de rééchantillonnage non paramétrique et paramétrique, où la probabilité d'observer une valeur supérieure à la plus grande donnée observée est calculée à partir de la distribution paramétrique.

2.1. MÉTHODES TRADITIONNELLES D'ESTIMATION DES PARAMÈTRES

2.1.1. Présentation du modèle de mélange de deux lois

Définissons d'abord la fonction de densité d'une variable aléatoire unidimensionnelle Y suivant un mélange de deux lois par:

$$f(y|\boldsymbol{\psi}) = (1 - \epsilon)f_1(y|\boldsymbol{\theta}_1) + \epsilon f_2(y|\boldsymbol{\theta}_2), \quad 0 \leq \epsilon \leq 1, \quad (2.1.1)$$

où $\boldsymbol{\psi} = (\epsilon, \boldsymbol{\theta}_1, \boldsymbol{\theta}_2)$, $\boldsymbol{\theta}_1$ et $\boldsymbol{\theta}_2$ représentent respectivement l'ensemble des paramètres des fonctions de densité $f(\cdot)$, $f_1(\cdot)$ et $f_2(\cdot)$. La fonction de densité $f(y|\boldsymbol{\psi})$ est un mélange de deux densités $f_1(\cdot)$ et $f_2(\cdot)$, respectivement pondérées par $1 - \epsilon$ et ϵ . Afin d'ajuster le modèle, nous devons choisir $f_1(\cdot)$ et $f_2(\cdot)$, puis estimer le vecteur des paramètres $\boldsymbol{\psi} = (\epsilon, \boldsymbol{\theta}_1, \boldsymbol{\theta}_2)$ par $\hat{\boldsymbol{\psi}} = (\hat{\epsilon}, \hat{\boldsymbol{\theta}}_1, \hat{\boldsymbol{\theta}}_2)$. Calculons l'espérance et la variance de Y :

THÉORÈME 2.1.1.

$$\begin{aligned} \mathbb{E}(Y) &= (1 - \epsilon)\mu_1 + \epsilon\mu_2 \\ &= \bar{\mu}, \end{aligned}$$

où μ_1 et μ_2 sont les espérances correspondant respectivement aux distributions $f_1(\cdot)$ et $f_2(\cdot)$.

$$\text{Var}(Y) = (1 - \epsilon) [\sigma_1^2 + (\mu_1 - \bar{\mu})^2] + \epsilon [\sigma_2^2 + (\mu_2 - \bar{\mu})^2],$$

où σ_1^2 et σ_2^2 sont les variances correspondant respectivement aux distributions $f_1(\cdot)$ et $f_2(\cdot)$.

DÉMONSTRATION

$$\begin{aligned} \mathbb{E}(Y) &= \int_y y f(y|\boldsymbol{\psi}) dy \\ &= \int_y y [(1 - \epsilon) f_1(y|\boldsymbol{\theta}_1) + \epsilon f_2(y|\boldsymbol{\theta}_2)] dy \\ &= (1 - \epsilon) \int_y y f_1(y|\boldsymbol{\theta}_1) dy + \epsilon \int_y y f_2(y|\boldsymbol{\theta}_2) dy \\ &= (1 - \epsilon) \mu_1 + \epsilon \mu_2 \\ &= \bar{\mu}, \end{aligned}$$

$$\begin{aligned} \text{Var}(Y) &= \int_y (y - \bar{\mu})^2 f(y|\boldsymbol{\psi}) dy \\ &= \int_y (y - \bar{\mu})^2 [(1 - \epsilon) f_1(y|\boldsymbol{\theta}_1) + \epsilon f_2(y|\boldsymbol{\theta}_2)] dy \\ &= (1 - \epsilon) \int_y (y - \bar{\mu})^2 f_1(y|\boldsymbol{\theta}_1) dy + \epsilon \int_y (y - \bar{\mu})^2 f_2(y|\boldsymbol{\theta}_2) dy \\ &= (1 - \epsilon) \int_y [(y - \mu_1) + (\mu_1 - \bar{\mu})]^2 f_1(y|\boldsymbol{\theta}_1) dy \\ &\quad + \epsilon \int_y [(y - \mu_2) + (\mu_2 - \bar{\mu})]^2 f_2(y|\boldsymbol{\theta}_2) dy \\ &= (1 - \epsilon) \int_y (y - \mu_1)^2 f_1(y|\boldsymbol{\theta}_1) dy + (1 - \epsilon) \int_y (\mu_1 - \bar{\mu})^2 f_1(y|\boldsymbol{\theta}_1) dy \\ &\quad + \epsilon \int_y (y - \mu_2)^2 f_2(y|\boldsymbol{\theta}_2) dy + \epsilon \int_y (\mu_2 - \bar{\mu})^2 f_2(y|\boldsymbol{\theta}_2) dy \\ &= (1 - \epsilon) [\sigma_1^2 + (\mu_1 - \bar{\mu})^2] + \epsilon [\sigma_2^2 + (\mu_2 - \bar{\mu})^2]. \end{aligned}$$

C.Q.F.D.

2.1.2. Méthode graphique

Plusieurs méthodes existent pour ajuster un mélange de lois. La plus facile est la méthode graphique qui permet de détecter visuellement le nombre de modes, leur moyenne et leur variance. C'est une bonne méthode pour produire des estimés initiaux des paramètres. Pour détecter les modes, nous pouvons dessiner l'histogramme des données pour avoir un aperçu de la fonction de densité ou encore dessiner la fonction de répartition empirique. Plusieurs techniques existent pour estimer les paramètres à partir de ces deux graphiques et elles sont décrites par Titterington, Smith et Makov (1985, section 4.1).

2.1.3. Méthode des moments

La méthode des moments, décrite par Cohen (1967), consiste à évaluer m moments théoriques aux m moments échantillonnaires correspondants. Ainsi, si nous avons m paramètres à estimer dans le modèle, il s'agit de créer un système de m équations impliquant les m paramètres. Chacune des équations est de la forme:

$$\mathbb{E}[Y^p] = \frac{1}{n} \sum_{i=1}^n y_i^p, p = 1, 2, \dots,$$

où n est la taille de l'échantillon. Appliquons cette méthode au cas particulier où $f_1(y|\theta_1)$ et $f_2(y|\theta_2)$ sont des densités normales, c'est-à-dire $\theta_1 = (\mu_1, \sigma_1^2)$, $\theta_2 = (\mu_2, \sigma_2^2)$ et donc $\psi = (\epsilon, \mu_1, \sigma_1^2, \mu_2, \sigma_2^2)$. Nous aurons d'abord besoin du résultat suivant que nous trouvons entre autres dans Bickel et Doksum (1977, section 1.8).

LEMME 2.1.1. *Si $Y \sim N(\mu, \sigma^2)$ alors le r -ième moment central de Y est*

$$\mathbb{E}(Y - \mu)^r = \begin{cases} \frac{r! \sigma^r}{2^{r/2} (r/2)!} & ; \text{ si } r \text{ est pair,} \\ 0 & ; \text{ si } r \text{ est impair.} \end{cases}$$

Nous avons donc

$$\begin{aligned}
\mathbb{E}[Y^p] &= \int_y y^p [(1 - \epsilon) f_1(y|\mu_1, \sigma_1^2) + \epsilon f_2(y|\mu_2, \sigma_2^2)] dy \\
&= (1 - \epsilon) \mathbb{E}_{f_1(y|\mu_1, \sigma_1^2)}[Y^p] + \epsilon \mathbb{E}_{f_2(y|\mu_2, \sigma_2^2)}[Y^p] \\
&= (1 - \epsilon) \mathbb{E}[(\sigma_1 Z + \mu_1)^p] + \epsilon \mathbb{E}[(\sigma_2 Z + \mu_2)^p],
\end{aligned}$$

où $Z \sim N(0, 1)$. Maintenant, en utilisant la formule du binôme de Newton, nous obtenons le théorème suivant:

THÉORÈME 2.1.2. *Pour p entier,*

$$\mathbb{E}[Y^p] = \sum_{r=0}^{\lfloor p/2 \rfloor} \binom{p}{2r} \frac{(2r)!}{2^r r!} [(1 - \epsilon) \sigma_1^{2r} \mu_1^{p-2r} + \epsilon \sigma_2^{2r} \mu_2^{p-2r}]$$

où $\lfloor p/2 \rfloor$ est la partie entière de $p/2$.

DÉMONSTRATION

$$\begin{aligned}
\mathbb{E}[Y^p] &= (1 - \epsilon) \mathbb{E} \left[\sum_{r=0}^p \binom{p}{r} (\sigma_1 Z)^r \mu_1^{p-r} \right] + \epsilon \mathbb{E} \left[\sum_{r=0}^p \binom{p}{r} (\sigma_2 Z)^r \mu_2^{p-r} \right] \\
&= \sum_{r=0}^p \binom{p}{r} \mathbb{E}[Z^r] [(1 - \epsilon) \sigma_1^r \mu_1^{p-r} + \epsilon \sigma_2^r \mu_2^{p-r}] \\
&= \sum_{r=0}^{\lfloor p/2 \rfloor} \binom{p}{2r} \mathbb{E}[Z^{2r}] [(1 - \epsilon) \sigma_1^{2r} \mu_1^{p-2r} + \epsilon \sigma_2^{2r} \mu_2^{p-2r}] \\
&= \sum_{r=0}^{\lfloor p/2 \rfloor} \binom{p}{2r} \frac{(2r)!}{2^r r!} [(1 - \epsilon) \sigma_1^{2r} \mu_1^{p-2r} + \epsilon \sigma_2^{2r} \mu_2^{p-2r}]
\end{aligned}$$

C.Q.F.D.

Pour notre problème, nous devons estimer $\epsilon, \mu_1, \sigma_1^2, \mu_2$ et σ_2^2 . Ainsi, nous devons résoudre le système non linéaire suivant de cinq équations à cinq inconnues:

$$\begin{aligned}\sum_{i=1}^n y_i/n &= (1 - \epsilon)\mu_1 + \epsilon\mu_2, \\ \sum_{i=1}^n y_i^2/n &= (1 - \epsilon)[\mu_1^2 + \sigma_1^2] + \epsilon[\mu_2^2 + \sigma_2^2], \\ \sum_{i=1}^n y_i^3/n &= (1 - \epsilon)[\mu_1^3 + 3\sigma_1^2\mu_1] + \epsilon[\mu_2^3 + 3\sigma_2^2\mu_2], \\ \sum_{i=1}^n y_i^4/n &= (1 - \epsilon)[\mu_1^4 + 6\sigma_1^2\mu_1^2 + 3\sigma_1^4] + \epsilon[\mu_2^4 + 6\sigma_2^2\mu_2^2 + 3\sigma_2^4], \\ \sum_{i=1}^n y_i^5/n &= (1 - \epsilon)[\mu_1^5 + 10\sigma_1^2\mu_1^3 + 15\sigma_1^4\mu_1] + \epsilon[\mu_2^5 + 10\sigma_2^2\mu_2^3 + 15\sigma_2^4\mu_2].\end{aligned}$$

Pour résoudre un tel système, nous pouvons utiliser la méthode de Newton telle que décrite dans Cohen (1967). Posons

$$f_p(\boldsymbol{\psi}) = \mathbb{E}[Y^p] - \frac{\sum_{i=1}^n y_i^p}{n} \text{ pour } p = 1, \dots, 5,$$

où $\boldsymbol{\psi} = (\psi_1, \psi_2, \psi_3, \psi_4, \psi_5) = (\epsilon, \mu_1, \sigma_1^2, \mu_2, \sigma_2^2)$. Nous cherchons donc $\hat{\boldsymbol{\psi}}$ qui vérifie $\mathbf{f}(\hat{\boldsymbol{\psi}}) = \mathbf{0}$ où $\mathbf{f}(\hat{\boldsymbol{\psi}}) = (f_1(\hat{\boldsymbol{\psi}}), f_2(\hat{\boldsymbol{\psi}}), f_3(\hat{\boldsymbol{\psi}}), f_4(\hat{\boldsymbol{\psi}}), f_5(\hat{\boldsymbol{\psi}}))$. La méthode de Newton est itérative. Elle consiste à approcher chacune des fonctions $f_p(\boldsymbol{\psi})$ par une fonction linéaire au voisinage d'un estimé préliminaire. Si nous notons l'estimé préliminaire par $\boldsymbol{\psi}^{(s)}$, alors les approximations linéaires sont:

$$f_p(\boldsymbol{\psi}) \approx f_p(\boldsymbol{\psi}^{(s)}) + \sum_{q=1}^5 \left[\frac{\partial f_p(\boldsymbol{\psi})}{\partial \psi_q} \Big|_{\boldsymbol{\psi}^{(s)}} (\psi_q - \psi_q^{(s)}) \right], \text{ pour } p = 1, \dots, 5,$$

à la condition que $f_p(\boldsymbol{\psi})$ soit deux fois continûment dérivable. Si nous définissons la matrice 5×5 suivante:

$$G(\boldsymbol{\psi}^{(s)}) = \{G_{p,q}(\boldsymbol{\psi}^{(s)})\} = \left\{ \frac{\partial f_p(\boldsymbol{\psi})}{\partial \psi_q} \Big|_{\boldsymbol{\psi}^{(s)}} \right\}$$

pour p et $q = 1, \dots, 5$, alors nous pouvons récrire les cinq équations à cinq inconnues sous la forme matricielle suivante:

$$\mathbf{f}(\boldsymbol{\psi}) \approx \mathbf{f}(\boldsymbol{\psi}^{(s)}) + G(\boldsymbol{\psi}^{(s)}) (\boldsymbol{\psi} - \boldsymbol{\psi}^{(s)}).$$

Puisque nous voulons trouver ψ qui vérifie $f(\psi) = 0$, nous trouvons alors

$$\psi = \psi^{(s+1)} = \psi^{(s)} - G^{-1}(\psi^{(s)}) f(\psi^{(s)})$$

si $G(\psi^{(s)})$ est inversible. Il s'agit donc de choisir des estimés initiaux $\psi^{(0)}$, par la méthode graphique par exemple. Si ces estimés sont trop loin de la solution, il est possible que le processus itératif ne converge pas. Aussi, s'il existe plusieurs racines, la racine vers laquelle la suite va converger dépendra des estimés initiaux. Il faut de plus choisir un critère de convergence absolue ou relative. Par exemple, nous pourrions arrêter le processus itératif lorsque $\max_{q=1,\dots,5} |\psi_q^{(s+1)} - \psi_q^{(s)}| < \delta_1$ où δ_1 est l'erreur absolue maximale choisie. Cependant, il est préférable d'arrêter le processus lorsque $\max_{q=1,\dots,5} \left(\max \left[\frac{\psi_q^{(s+1)}}{\psi_q^{(s)}}, \frac{\psi_q^{(s)}}{\psi_q^{(s+1)}} \right] \right) < 1 + \delta_2$, où δ_2 est l'erreur relative maximale choisie.

Si nous calculons $G(\psi^{(s)})$ pour notre exemple de mélange de lois normales, nous trouvons:

$$\begin{aligned} \frac{\partial f}{\partial \epsilon} &= \begin{pmatrix} \mu_2 - \mu_1 \\ (\mu_2^2 + \sigma_2^2) - (\mu_1^2 + \sigma_1^2) \\ (\mu_2^3 + 3\sigma_2^2\mu_2) - (\mu_1^3 + 3\sigma_1^2\mu_1) \\ (\mu_2^4 + 6\sigma_2^2\mu_2^2 + 3\sigma_2^4) - (\mu_1^4 + 6\sigma_1^2\mu_1^2 + 3\sigma_1^4) \\ (\mu_2^5 + 10\sigma_2^2\mu_2^3 + 15\sigma_2^4\mu_2) - (\mu_1^5 + 10\sigma_1^2\mu_1^3 + 15\sigma_1^4\mu_1) \end{pmatrix}, \\ \frac{\partial f}{\partial \mu_2} &= \begin{pmatrix} \epsilon \\ \epsilon(2\mu_2) \\ \epsilon(3\mu_2^2 + 3\sigma_2^2) \\ \epsilon(4\mu_2^3 + 12\sigma_2^2\mu_2) \\ \epsilon(5\mu_2^4 + 30\sigma_2^2\mu_2^2 + 15\sigma_2^4) \end{pmatrix}, \\ \frac{\partial f}{\partial \sigma_2^2} &= \begin{pmatrix} 0 \\ \epsilon \\ \epsilon(3\mu_2) \\ \epsilon(6\mu_2^2 + 6\sigma_2^2) \\ \epsilon(10\mu_2^3 + 30\sigma_2^2\mu_2) \end{pmatrix}. \end{aligned}$$

Nous trouvons la même chose pour les dérivées par rapport à μ_1 et σ_1^2 en remplaçant ϵ par $1 - \epsilon$ et l'indice 2 par 1 dans les dérivées par rapport à μ_2 et σ_2^2 .

2.1.4. Méthode du maximum de vraisemblance

Comme son nom l'indique, cette méthode, décrite par Rao (1948), consiste à trouver les estimateurs de $\boldsymbol{\psi}$ qui maximisent la fonction de vraisemblance, notée

$$L(\boldsymbol{\psi}|\mathbf{y}) = \prod_{i=1}^n [(1 - \epsilon) f_1(y_i|\boldsymbol{\theta}_1) + \epsilon f_2(y_i|\boldsymbol{\theta}_2)].$$

Nous pouvons chercher plutôt à maximiser le logarithme de la fonction de vraisemblance, ce qui est équivalent et généralement plus facile. Nous maximiserons donc

$$l(\boldsymbol{\psi}|\mathbf{y}) = \log L(\boldsymbol{\psi}|\mathbf{y}) = \sum_{i=1}^n \log [(1 - \epsilon) f_1(y_i|\boldsymbol{\theta}_1) + \epsilon f_2(y_i|\boldsymbol{\theta}_2)].$$

La valeur de $\boldsymbol{\psi}$ qui maximise cette fonction est celle qui satisfait

$$\frac{\partial l(\boldsymbol{\psi}|\mathbf{y})}{\partial \boldsymbol{\psi}} = \mathbf{0},$$

si le maximum de $l(\boldsymbol{\psi}|\mathbf{y})$ n'est pas situé aux extrêmes des valeurs possibles de $\boldsymbol{\psi}$. Dans les problèmes d'estimation des paramètres d'un mélange de lois, les équations sont non linéaires en $\boldsymbol{\psi}$ et alors aucune forme explicite ne peut être trouvée pour les estimateurs. La solution est obtenue par une méthode itérative, comme la méthode de Newton vue à la section 2.1.3.

Toutefois un algorithme beaucoup plus simple à appliquer pour la méthode du maximum de vraisemblance est l'algorithme EM (Espérance-Maximisation), qui a d'abord été introduit par Dempster, Laird et Rubin (1977) et qui aussi est bien décrit dans Titterington, Smith et Makov (1985, section 4.3).

2.2. ALGORITHME EM

2.2.1. Description de l'algorithme EM

Supposons que nous voulions trouver l'estimateur $\hat{\boldsymbol{\psi}}$ de $\boldsymbol{\psi}$ par la méthode du maximum de vraisemblance et que la fonction de densité soit donnée par $f(y|\boldsymbol{\psi})$ tel que définie par l'équation (2.1.1). Comme nous venons de le voir à la section précédente, il s'agit alors de maximiser par rapport à $\boldsymbol{\psi}$ le logarithme de

la fonction de vraisemblance, soit

$$l(\boldsymbol{\psi}|\mathbf{y}) = \log L(\boldsymbol{\psi}|\mathbf{y}) = \log f(\mathbf{y}|\boldsymbol{\psi}) = \log \prod_{i=1}^n f(y_i|\boldsymbol{\psi}) = \sum_{i=1}^n \log f(y_i|\boldsymbol{\psi}),$$

en résolvant l'équation

$$\frac{\partial l(\boldsymbol{\psi}|\mathbf{y})}{\partial \boldsymbol{\psi}} = \mathbf{0}.$$

Il est parfois très difficile de maximiser cette équation, comme dans la section précédente où nous cherchions à estimer les paramètres d'un mélange de lois. Nous pouvons simplifier les calculs en introduisant des variables dites latentes, que nous dénotons par $\mathbf{z} = (z_1, \dots, z_n)$. Par exemple, dans le cas de l'estimation de paramètres d'un mélange de deux lois, nous pouvons introduire z_i comme une variable ayant une distribution $Bin(1, \epsilon)$ nous indiquant de quelle loi provient y_i . Les données complètes sont alors $(\mathbf{y}, \mathbf{z}) = (y_1, \dots, y_n, z_1, \dots, z_n)$. Notons que y_i est indépendante de $y_j, \forall i \neq j, i, j = 1, \dots, n$ et de la même façon z_k est indépendante de $z_l, \forall k \neq l, k, l = 1, \dots, n$. Si nous voulons trouver $\hat{\boldsymbol{\psi}}$ comme ci-haut en maximisant le logarithme de la fonction de vraisemblance $l(\boldsymbol{\psi}|\mathbf{y}, \mathbf{z})$, nous ne pouvons le faire directement sans connaître \mathbf{z} . Nous pourrions réussir à maximiser $l(\boldsymbol{\psi}|\mathbf{y}, \mathbf{z})$ par itération à l'aide de l'algorithme EM.

Tout d'abord, nous devons identifier les fonctions de densité de $\mathbf{z}|\boldsymbol{\psi}$ et $\mathbf{y}|\mathbf{z}, \boldsymbol{\psi}$. Nous trouvons par la suite le logarithme de la fonction de vraisemblance:

$$l(\boldsymbol{\psi}|\mathbf{y}, \mathbf{z}) = \log L(\boldsymbol{\psi}|\mathbf{y}, \mathbf{z}) = \log f(\mathbf{y}, \mathbf{z}|\boldsymbol{\psi}) = \log [f(\mathbf{y}|\mathbf{z}, \boldsymbol{\psi}) f(\mathbf{z}|\boldsymbol{\psi})].$$

L'algorithme EM consiste à générer une suite d'estimateurs $\{\boldsymbol{\psi}^{(m)}\}$ à partir d'un estimateur initial $\{\boldsymbol{\psi}^{(0)}\}$ jusqu'à convergence vers $\hat{\boldsymbol{\psi}}$. Chaque itération notée m consiste en deux étapes:

1) Étape E (Espérance)

Calculer

$$Q(\boldsymbol{\psi}, \boldsymbol{\psi}^{(m)}) = \mathbb{E}^{\mathbf{z}|\boldsymbol{\psi}^{(m)}, \mathbf{y}} [l(\boldsymbol{\psi}|\mathbf{y}, \mathbf{z})] = \mathbb{E}^{\mathbf{z}|\boldsymbol{\psi}^{(m)}, \mathbf{y}} [\log [f(\mathbf{y}|\mathbf{z}, \boldsymbol{\psi}) f(\mathbf{z}|\boldsymbol{\psi})]].$$

Nous trouvons la distribution de $\mathbf{z}|\boldsymbol{\psi}^{(m)}, \mathbf{y}$ par la relation suivante:

$$f(\mathbf{z}|\boldsymbol{\psi}^{(m)}, \mathbf{y}) = \frac{f(\mathbf{y}, \mathbf{z}|\boldsymbol{\psi}^{(m)})}{f(\mathbf{y}|\boldsymbol{\psi}^{(m)})} \propto f(\mathbf{y}, \mathbf{z}|\boldsymbol{\psi}^{(m)}).$$

2) Étape M (Maximisation)

Trouver $\psi = \psi^{(m+1)}$ qui maximise $Q(\psi, \psi^{(m)})$ en résolvant l'équation

$$\frac{\partial Q(\psi, \psi^{(m)})}{\partial \psi} = \mathbf{0}.$$

2.2.2. Application de l'algorithme EM à l'ajustement de distributions bimodales

Nous pouvons récrire la fonction de densité bimodale

$$f(y_i|\psi) = f(y_i|\epsilon, \theta_1, \theta_2) = (1 - \epsilon)f_1(y_i|\theta_1) + \epsilon f_2(y_i|\theta_2), \quad i = 1, \dots, n$$

en introduisant les variables latentes z_1, \dots, z_n comme suit:

$$f(y_i|\mathbf{z}, \psi) = f(y_i|z_i, \theta_1, \theta_2) = f_1(y_i|\theta_1)^{1-z_i} f_2(y_i|\theta_2)^{z_i}$$

où $z_i = z_i|\psi = z_i|\epsilon \sim \text{Bin}(1, \epsilon)$, $i = 1, \dots, n$.

Nous interprétons la variable latente z_i comme une variable indicatrice nous révélant de quelle distribution $f_1(\cdot)$ ou $f_2(\cdot)$ provient y_i . Puisque $z_i|\epsilon \sim \text{Bin}(1, \epsilon)$, z_i prendra la valeur 1 dans une proportion ϵ et alors nous aurons $f(y_i|z_i, \theta_1, \theta_2) = f_2(y_i|\theta_2)$. De la même façon, z_i prendra la valeur 0 dans une proportion $1 - \epsilon$ et alors nous aurons $f(y_i|z_i, \theta_1, \theta_2) = f_1(y_i|\theta_1)$. Évidemment nous ne connaissons pas z_i . Nous supposons aussi que y_i est indépendante de z_j , $\forall i \neq j$, $i, j = 1, \dots, n$. De plus, nous remarquons que $z_i|\epsilon$ ne dépend pas de θ_1 ni de θ_2 , et que $y_i|z_i, \theta_1, \theta_2$ ne dépend pas de ϵ .

Appliquons maintenant l'algorithme EM à ce modèle de mélange de lois pour estimer les paramètres ψ . Trouvons la distribution de $z_i|\psi, \mathbf{y}$ où $\psi = (\epsilon, \theta_1, \theta_2)$. Nous avons le théorème suivant:

THÉORÈME 2.2.1.

$$z_i|\theta_1, \theta_2, \epsilon, \mathbf{y} \sim \text{Bin} \left(1, \frac{\epsilon f_2(y_i|\theta_2)}{(1 - \epsilon) f_1(y_i|\theta_1) + \epsilon f_2(y_i|\theta_2)} \right).$$

DÉMONSTRATION

$$\begin{aligned}
f(z_i|\boldsymbol{\theta}_1, \boldsymbol{\theta}_2, \epsilon, \mathbf{y}) &= f(z_i|y_i, \boldsymbol{\theta}_1, \boldsymbol{\theta}_2, \epsilon) \\
&= \frac{f(y_i|z_i, \boldsymbol{\theta}_1, \boldsymbol{\theta}_2, \epsilon) f(z_i|\boldsymbol{\theta}_1, \boldsymbol{\theta}_2, \epsilon)}{f(y_i|\boldsymbol{\theta}_1, \boldsymbol{\theta}_2, \epsilon)} \\
&\propto f(y_i|z_i, \boldsymbol{\theta}_1, \boldsymbol{\theta}_2) f(z_i|\epsilon) \\
&= f_1(y_i|\boldsymbol{\theta}_1)^{1-z_i} f_2(y_i|\boldsymbol{\theta}_2)^{z_i} (1-\epsilon)^{1-z_i} \epsilon^{z_i} \\
&\propto \left[\frac{(1-\epsilon) f_1(y_i|\boldsymbol{\theta}_1)}{(1-\epsilon) f_1(y_i|\boldsymbol{\theta}_1) + \epsilon f_2(y_i|\boldsymbol{\theta}_2)} \right]^{1-z_i} \\
&\quad \times \left[\frac{\epsilon f_2(y_i|\boldsymbol{\theta}_2)}{(1-\epsilon) f_1(y_i|\boldsymbol{\theta}_1) + \epsilon f_2(y_i|\boldsymbol{\theta}_2)} \right]^{z_i} \\
\Rightarrow z_i|\boldsymbol{\theta}_1, \boldsymbol{\theta}_2, \epsilon, \mathbf{y} &\sim B\left(1, \frac{\epsilon f_2(y_i|\boldsymbol{\theta}_2)}{(1-\epsilon) f_1(y_i|\boldsymbol{\theta}_1) + \epsilon f_2(y_i|\boldsymbol{\theta}_2)}\right).
\end{aligned}$$

C.Q.F.D.

Effectuons l'étape E, c'est-à-dire calculer $Q(\boldsymbol{\psi}, \boldsymbol{\psi}^{(m)})$. Nous avons le théorème suivant:

THÉORÈME 2.2.2.

$$\begin{aligned}
Q(\boldsymbol{\psi}, \boldsymbol{\psi}^{(m)}) &= \sum_{i=1}^n [1 - \mathbb{E}^{\mathbf{z}|\boldsymbol{\psi}^{(m)}, \mathbf{y}}(z_i)] \log f_1(y_i|\boldsymbol{\theta}_1) \\
&\quad + \sum_{i=1}^n \mathbb{E}^{\mathbf{z}|\boldsymbol{\psi}^{(m)}, \mathbf{y}}(z_i) \log f_2(y_i|\boldsymbol{\theta}_2) + \log(1-\epsilon) \sum_{i=1}^n [1 - \mathbb{E}^{\mathbf{z}|\boldsymbol{\psi}^{(m)}, \mathbf{y}}(z_i)] \\
&\quad + \log \epsilon \sum_{i=1}^n \mathbb{E}^{\mathbf{z}|\boldsymbol{\psi}^{(m)}, \mathbf{y}}(z_i)
\end{aligned}$$

où

$$\begin{aligned}
\mathbb{E}^{\mathbf{z}|\boldsymbol{\psi}^{(m)}, \mathbf{y}}(z_i) &= \mathbb{E}^{z_i|\boldsymbol{\psi}^{(m)}, y_i}(z_i) \\
&= \frac{\epsilon^{(m)} f_2(y_i|\boldsymbol{\theta}_2^{(m)})}{(1-\epsilon^{(m)}) f_1(y_i|\boldsymbol{\theta}_1^{(m)}) + \epsilon^{(m)} f_2(y_i|\boldsymbol{\theta}_2^{(m)})}.
\end{aligned}$$

DÉMONSTRATION

$$\begin{aligned}
Q(\boldsymbol{\psi}, \boldsymbol{\psi}^{(m)}) &= \mathbb{E}^{\mathbf{z}|\boldsymbol{\psi}^{(m)}, \mathbf{y}} [\log[f(\mathbf{y}|\mathbf{z}, \boldsymbol{\psi})f(\mathbf{z}|\boldsymbol{\psi})]] \\
&= \mathbb{E}^{\mathbf{z}|\boldsymbol{\psi}^{(m)}, \mathbf{y}} \left[\log \left[\prod_{i=1}^n [f(y_i|z_i, \boldsymbol{\theta}_1, \boldsymbol{\theta}_2) f(z_i|\epsilon)] \right] \right] \\
&= \mathbb{E}^{\mathbf{z}|\boldsymbol{\psi}^{(m)}, \mathbf{y}} \left[\sum_{i=1}^n [(1 - z_i) \log f_1(y_i|\boldsymbol{\theta}_1) + z_i \log f_2(y_i|\boldsymbol{\theta}_2) \right. \\
&\quad \left. + (1 - z_i) \log(1 - \epsilon) + z_i \log \epsilon] \right] \\
&= \sum_{i=1}^n [1 - \mathbb{E}^{\mathbf{z}|\boldsymbol{\psi}^{(m)}, \mathbf{y}}(z_i)] \log f_1(y_i|\boldsymbol{\theta}_1) \\
&\quad + \sum_{i=1}^n \mathbb{E}^{\mathbf{z}|\boldsymbol{\psi}^{(m)}, \mathbf{y}}(z_i) \log f_2(y_i|\boldsymbol{\theta}_2) + \log(1 - \epsilon) \sum_{i=1}^n [1 - \mathbb{E}^{\mathbf{z}|\boldsymbol{\psi}^{(m)}, \mathbf{y}}(z_i)] \\
&\quad + \log \epsilon \sum_{i=1}^n \mathbb{E}^{\mathbf{z}|\boldsymbol{\psi}^{(m)}, \mathbf{y}}(z_i)
\end{aligned}$$

C.Q.F.D.

L'Étape M consiste à maximiser $Q(\boldsymbol{\psi}, \boldsymbol{\psi}^{(m)})$ par rapport à $\boldsymbol{\psi} = (\epsilon, \boldsymbol{\theta}_1, \boldsymbol{\theta}_2)$. Nous avons le théorème suivant:

THÉORÈME 2.2.3.

$$\epsilon^{(m+1)} = \frac{\sum_{i=1}^n \mathbb{E}^{z_i|\boldsymbol{\psi}^{(m)}, y_i}(z_i)}{n}.$$

DÉMONSTRATION

Nous trouvons $\epsilon^{(m+1)}$ en résolvant l'équation suivante pour ϵ :

$$\begin{aligned}
\frac{\partial Q(\boldsymbol{\psi}, \boldsymbol{\psi}^{(m)})}{\partial \epsilon} &= \frac{\partial}{\partial \epsilon} \left[\log(1 - \epsilon) \sum_{i=1}^n [1 - \mathbb{E}^{z_i|\boldsymbol{\psi}^{(m)}, y_i}(z_i)] + \log \epsilon \sum_{i=1}^n \mathbb{E}^{z_i|\boldsymbol{\psi}^{(m)}, y_i}(z_i) \right] \\
&= \frac{\sum_{i=1}^n \mathbb{E}^{z_i|\boldsymbol{\psi}^{(m)}, y_i}(z_i)}{\epsilon} - \frac{\sum_{i=1}^n [1 - \mathbb{E}^{z_i|\boldsymbol{\psi}^{(m)}, y_i}(z_i)]}{(1 - \epsilon)} = 0 \\
\Rightarrow \epsilon^{(m+1)} &= \frac{\sum_{i=1}^n \mathbb{E}^{z_i|\boldsymbol{\psi}^{(m)}, y_i}(z_i)}{n}.
\end{aligned}$$

C.Q.F.D.

Nous trouvons $\boldsymbol{\theta}_1^{(m+1)}$ en résolvant l'équation suivante pour $\boldsymbol{\theta}_1$:

$$\frac{\partial Q(\boldsymbol{\psi}, \boldsymbol{\psi}^{(m)})}{\partial \boldsymbol{\theta}_1} = \sum_{i=1}^n [1 - \mathbb{E}^{z_i|\boldsymbol{\psi}^{(m)}, y_i}(z_i)] \frac{\partial}{\partial \boldsymbol{\theta}_1} [\log f_1(y_i|\boldsymbol{\theta}_1)] = 0.$$

De la même façon, nous trouvons $\theta_2^{(m+1)}$ en résolvant l'équation suivante pour θ_2 :

$$\frac{\partial Q(\psi, \psi^{(m)})}{\partial \theta_2} = \sum_{i=1}^n \mathbb{E}^{z_i | \psi^{(m)}, y_i}(z_i) \frac{\partial}{\partial \theta_2} [\log f_2(y_i | \theta_2)] = 0.$$

2.2.3. Application de l'algorithme EM à des mélanges de lois spécifiques

Appliquons maintenant l'algorithme EM à des mélanges de lois spécifiques. Nous avons choisi les mélanges de lois normales, gamma, log-normales et inverse-gamma. Ces quatre distributions nous permettront de modéliser adéquatement une grande variété de types de données.

2.2.3.1. Mélanges de lois normales

La loi normale présente l'avantage d'être la plus connue des distributions et d'être relativement facile à utiliser. De plus, par le théorème limite central, la distribution de la moyenne empirique converge vers une loi normale, quelle que soit la loi des observations, lorsque la taille de l'échantillon tend vers l'infini. Il est donc naturel de considérer d'abord un mélange de lois normales.

Si $y_i | \theta_1 \sim N(\mu_1, \sigma_1^2)$ et $y_i | \theta_2 \sim N(\mu_2, \sigma_2^2)$, alors pour $(\mu, \sigma^2) = (\mu_1, \sigma_1^2)$ ou $(\mu, \sigma^2) = (\mu_2, \sigma_2^2)$, nous avons

$$f(y_i | \theta) = f(y_i | \mu, \sigma^2) = (2\pi\sigma^2)^{-1/2} e^{-\frac{1}{2\sigma^2}(y_i - \mu)^2}, \quad -\infty < y_i, \mu < \infty, \quad 0 < \sigma^2 < \infty,$$

et

$$\log f(y_i | \mu, \sigma^2) = -\frac{\log(2\pi)}{2} - \frac{\log(\sigma^2)}{2} - \frac{(y_i - \mu)^2}{2\sigma^2}.$$

À partir d'estimés préliminaires $\boldsymbol{\psi}^{(m)} = (\epsilon^{(m)}, \mu_1^{(m)}, \sigma_1^{2(m)}, \mu_2^{(m)}, \sigma_2^{2(m)})$, nous trouvons donc à l'étape E:

$$\begin{aligned} \mathbb{E}^{z|\boldsymbol{\psi}^{(m)}, \mathbf{y}}(z_i) &= \frac{\epsilon^{(m)} f_2(y_i|\boldsymbol{\theta}_2^{(m)})}{(1 - \epsilon^{(m)}) f_1(y_i|\boldsymbol{\theta}_1^{(m)}) + \epsilon^{(m)} f_2(y_i|\boldsymbol{\theta}_2^{(m)})} \\ &= \epsilon^{(m)} (2\pi\sigma_2^{2(m)})^{-1/2} \exp\left[-\frac{1}{2\sigma_2^{2(m)}}(y_i - \mu_2^{(m)})^2\right] \\ &\times \left((1 - \epsilon^{(m)}) (2\pi\sigma_1^{2(m)})^{-1/2} \exp\left[-\frac{1}{2\sigma_1^{2(m)}}(y_i - \mu_1^{(m)})^2\right] \right. \\ &\left. + \epsilon^{(m)} (2\pi\sigma_2^{2(m)})^{-1/2} \exp\left[-\frac{1}{2\sigma_2^{2(m)}}(y_i - \mu_2^{(m)})^2\right] \right)^{-1}. \end{aligned}$$

Nous obtenons ensuite à l'étape M:

$$\epsilon^{(m+1)} = \frac{\sum_{i=1}^n \mathbb{E}^{z_i|\boldsymbol{\psi}^{(m)}, y_i}(z_i)}{n}.$$

Nous calculons $\mu_2^{(m+1)}$ en résolvant l'équation suivante:

$$0 = \sum_{i=1}^n \mathbb{E}^{z_i|\boldsymbol{\psi}^{(m)}, y_i}(z_i) \frac{\partial}{\partial \mu_2} \left[\log f_2(y_i|\mu_2, \sigma_2^2) \right] = - \sum_{i=1}^n \mathbb{E}^{z_i|\boldsymbol{\psi}^{(m)}, y_i}(z_i) \frac{(y_i - \mu_2)}{\sigma_2^2}.$$

Nous trouvons alors

$$\mu_2^{(m+1)} = \frac{\sum_{i=1}^n y_i \mathbb{E}^{z_i|\boldsymbol{\psi}^{(m)}, y_i}(z_i)}{\sum_{i=1}^n \mathbb{E}^{z_i|\boldsymbol{\psi}^{(m)}, y_i}(z_i)}.$$

De la même façon, nous obtenons

$$\mu_1^{(m+1)} = \frac{\sum_{i=1}^n y_i [1 - \mathbb{E}^{z_i|\boldsymbol{\psi}^{(m)}, y_i}(z_i)]}{\sum_{i=1}^n [1 - \mathbb{E}^{z_i|\boldsymbol{\psi}^{(m)}, y_i}(z_i)]}.$$

Nous calculons aussi $\sigma_2^{2(m+1)}$ en résolvant l'équation suivante:

$$\begin{aligned} 0 &= \sum_{i=1}^n \mathbb{E}^{z_i|\boldsymbol{\psi}^{(m)}, y_i}(z_i) \frac{\partial}{\partial \sigma_2^2} \left[\log f_2(y_i|\mu_2, \sigma_2^2) \right] \\ &= \sum_{i=1}^n \mathbb{E}^{z_i|\boldsymbol{\psi}^{(m)}, y_i}(z_i) \frac{1}{2} \left[\left(\frac{y_i - \mu_2}{\sigma_2^2} \right)^2 - \left(\frac{1}{\sigma_2^2} \right) \right]. \end{aligned}$$

Nous trouvons alors

$$\sigma_2^{2(m+1)} = \frac{\sum_{i=1}^n (y_i - \mu_2^{(m+1)})^2 \mathbb{E}^{z_i|\boldsymbol{\psi}^{(m)}, y_i}(z_i)}{\sum_{i=1}^n \mathbb{E}^{z_i|\boldsymbol{\psi}^{(m)}, y_i}(z_i)}.$$

De la même façon, nous obtenons

$$\sigma_1^{2(m+1)} = \frac{\sum_{i=1}^n (y_i - \mu_1^{(m+1)})^2 (1 - \mathbb{E}^{z_i|\psi^{(m)}, y_i}(z_i))}{\sum_{i=1}^n (1 - \mathbb{E}^{z_i|\psi^{(m)}, y_i}(z_i))}.$$

2.2.3.2. Mélange de lois gamma

Si $y_i|\theta_1 \sim G(\alpha_1, \beta_1)$ et $y_i|\theta_2 \sim G(\alpha_2, \beta_2)$, alors pour $(\alpha, \beta) = (\alpha_1, \beta_1)$ ou $(\alpha, \beta) = (\alpha_2, \beta_2)$, nous avons

$$f(y_i|\theta) = f(y_i|\alpha, \beta) = \frac{\beta^\alpha}{\Gamma(\alpha)} y_i^{\alpha-1} e^{-\beta y_i}, y_i, \alpha, \beta > 0,$$

et

$$\log f(y_i|\alpha, \beta) = \alpha \log \beta - \log \Gamma(\alpha) + (\alpha - 1) \log y_i - \beta y_i.$$

À partir d'estimés préliminaires $\psi^{(m)} = (\epsilon^{(m)}, \alpha_1^{(m)}, \beta_1^{(m)}, \alpha_2^{(m)}, \beta_2^{(m)})$, nous trouvons donc à l'étape E:

$$\begin{aligned} \mathbb{E}^{z|\psi^{(m)}, y}(z_i) &= \frac{\epsilon^{(m)} f_2(y_i|\theta_2^{(m)})}{(1 - \epsilon^{(m)}) f_1(y_i|\theta_1^{(m)}) + \epsilon^{(m)} f_2(y_i|\theta_2^{(m)})} \\ &= \frac{\epsilon^{(m)} \frac{\beta_2^{\alpha_2}}{\Gamma(\alpha_2)} y_i^{\alpha_2-1} e^{-\beta_2 y_i}}{(1 - \epsilon^{(m)}) \frac{\beta_1^{\alpha_1}}{\Gamma(\alpha_1)} y_i^{\alpha_1-1} e^{-\beta_1 y_i} + \epsilon^{(m)} \frac{\beta_2^{\alpha_2}}{\Gamma(\alpha_2)} y_i^{\alpha_2-1} e^{-\beta_2 y_i}}. \end{aligned}$$

Nous obtenons ensuite à l'étape M:

$$\epsilon^{(m+1)} = \frac{\sum_{i=1}^n \mathbb{E}^{z_i|\psi^{(m)}, y_i}(z_i)}{n}.$$

Nous calculons $\alpha_2^{(m+1)}$ en résolvant l'équation suivante:

$$\begin{aligned} 0 &= \sum_{i=1}^n \mathbb{E}^{z_i|\psi^{(m)}, y_i}(z_i) \frac{\partial}{\partial \alpha_2} [\log f_2(y_i|\alpha_2, \beta_2)] \\ &= \sum_{i=1}^n \mathbb{E}^{z_i|\psi^{(m)}, y_i}(z_i) \left[\log \beta_2 - \frac{\Gamma'(\alpha_2)}{\Gamma(\alpha_2)} + \log y_i \right]. \end{aligned}$$

Nous trouvons alors

$$\frac{\Gamma'(\alpha_2^{(m+1)})}{\Gamma(\alpha_2^{(m+1)})} = \log \beta_2^{(m)} + \frac{\sum_{i=1}^n (\log y_i) \mathbb{E}^{z_i|\psi^{(m)}, y_i}(z_i)}{\sum_{i=1}^n \mathbb{E}^{z_i|\psi^{(m)}, y_i}(z_i)}.$$

De la même façon, nous obtenons

$$\frac{\Gamma'(\alpha_1^{(m+1)})}{\Gamma(\alpha_1^{(m+1)})} = \log \beta_1^{(m)} + \frac{\sum_{i=1}^n (\log y_i) (1 - \mathbb{E}^{z_i|\psi^{(m)}, y_i}(z_i))}{\sum_{i=1}^n (1 - \mathbb{E}^{z_i|\psi^{(m)}, y_i}(z_i))}.$$

La fonction $\frac{\Gamma'(\alpha)}{\Gamma(\alpha)}$ est appelée la fonction digamma (voir Wolfram (1991, section 3.2)) et nous pouvons la calculer par la formule suivante:

$$\frac{\Gamma'(\alpha)}{\Gamma(\alpha)} = -c - \frac{1}{\alpha} + \alpha \sum_{k=1}^{\infty} \frac{1}{k(\alpha+k)},$$

où $c=0,577215669901\dots$ est la constante d'Euler. Alors pour $\alpha_2^{(m+1)}$, nous avons l'équation suivante:

$$-c - \frac{1}{\alpha_2^{(m+1)}} + \alpha_2^{(m+1)} \sum_{k=1}^{\infty} \frac{1}{k(\alpha_2^{(m+1)} + k)} = \log \beta_2^{(m)} + \frac{\sum_{i=1}^n (\log y_i) \mathbb{E}^{z_i | \psi^{(m)}, y_i}(z_i)}{\sum_{i=1}^n \mathbb{E}^{z_i | \psi^{(m)}, y_i}(z_i)}$$

que nous pouvons résoudre pour $\alpha_2^{(m+1)}$ de façon itérative, soit:

$$\alpha_2^{(m+1)} = \frac{c + (1/\alpha_2^{(m)}) + \left(\log \beta_2^{(m)} + \frac{\sum_{i=1}^n (\log y_i) \mathbb{E}^{z_i | \psi^{(m)}, y_i}(z_i)}{\sum_{i=1}^n \mathbb{E}^{z_i | \psi^{(m)}, y_i}(z_i)} \right)}{\sum_{k=1}^{\infty} \left(k(\alpha_2^{(m)} + k) \right)^{-1}}.$$

Nous trouvons les mêmes résultats pour $\alpha_1^{(m+1)}$ en remplaçant l'indice 2 par 1 et $\mathbb{E}^{z_i | \psi^{(m)}, y_i}(z_i)$ par $1 - \mathbb{E}^{z_i | \psi^{(m)}, y_i}(z_i)$.

Nous calculons aussi $\beta_2^{(m+1)}$ en résolvant l'équation suivante:

$$\begin{aligned} 0 &= \sum_{i=1}^n \mathbb{E}^{z_i | \psi^{(m)}, y_i}(z_i) \frac{\partial}{\partial \beta_2} [\log f_2(y_i | \alpha_2, \beta_2)] \\ &= \sum_{i=1}^n \mathbb{E}^{z_i | \psi^{(m)}, y_i}(z_i) \left(\frac{\alpha_2}{\beta_2} - y_i \right). \end{aligned}$$

Nous obtenons alors

$$\beta_2^{(m+1)} = \frac{\alpha_2^{(m+1)} \sum_{i=1}^n \mathbb{E}^{z_i | \psi^{(m)}, y_i}(z_i)}{\sum_{i=1}^n y_i \mathbb{E}^{z_i | \psi^{(m)}, y_i}(z_i)}.$$

De la même façon, nous trouvons

$$\beta_1^{(m+1)} = \frac{\alpha_1^{(m+1)} \sum_{i=1}^n (1 - \mathbb{E}^{z_i | \psi^{(m)}, y_i}(z_i))}{\sum_{i=1}^n y_i (1 - \mathbb{E}^{z_i | \psi^{(m)}, y_i}(z_i))}.$$

2.2.3.3. Mélange de lois log-normales

Nous savons que $Y \sim LN(\mu, \sigma^2)$ si $\log Y \sim N(\mu, \sigma^2)$, $Y > 0$. Afin de trouver la fonction de densité d'une log-normale, effectuons le changement de variables suivant:

$$X = \log Y \text{ et } dX = dY/Y$$

où $X \sim N(\mu, \sigma^2)$. Alors

$$f_Y(y|\mu, \sigma^2) = \frac{1}{y} f_X(\log y|\mu, \sigma^2) = \frac{1}{y} \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2\sigma^2}(\log y - \mu)^2},$$

$$y > 0, -\infty < \mu < \infty, \sigma > 0.$$

La fonction de densité d'une variable aléatoire Y suivant une distribution log-normale est la fonction de densité d'une loi normale en remplaçant y par $\log y$ et en la divisant par y . Aussi nous pouvons voir que résoudre cette équation:

$$\frac{\partial}{\partial \theta} \log f_Y(y|\mu, \sigma^2) = \frac{\partial}{\partial \theta} \log \frac{1}{y} f_X(\log y|\mu, \sigma^2) = 0$$

pour $\theta = (\mu, \sigma^2)$ est équivalent à résoudre cette équation:

$$\frac{\partial}{\partial \theta} \log f_X(\log y|\mu, \sigma^2) = 0,$$

où

$$Y \sim LN(\mu, \sigma^2) \text{ et } X \sim N(\mu, \sigma^2).$$

Nous pouvons donc estimer les paramètres d'un mélange de lois log-normales de la même façon que nous le faisons pour un mélange de lois normales, en remplaçant simplement les données originales par leur logarithme.

2.2.3.4. Mélange de lois inverse-gamma

Nous savons que $Y \sim IG(\alpha, \beta)$ si $1/Y \sim G(\alpha, \beta)$, $Y > 0$. Afin de trouver la fonction de densité d'une inverse-gamma, faisons le changement de variables suivant:

$$X = 1/Y \text{ et } dX = -dY/Y^2$$

où $X \sim G(\alpha, \beta)$. Alors nous obtenons

$$f_Y(y|\alpha, \beta) = \frac{1}{y^2} f_X(y^{-1}|\alpha, \beta) = \frac{\beta^\alpha}{\Gamma(\alpha)y^{\alpha+1}} e^{-\beta/y},$$

$$y > 0, \alpha, \beta > 0.$$

La fonction de densité d'une variable aléatoire Y suivant une distribution inverse-gamma est la fonction de densité d'une loi gamma en remplaçant y par $1/y$ et en

TABLEAU 2.1. Ajustement de distributions bimodales pour le traitement à la TMP/SMX

| | normale | log-normale (EM) | log-normale (moments) | gamma | inverse-gamma |
|--------------------------|---------|---------------------|--------------------------|-------|---------------|
| ϵ | 0,178 | 0,162 | 0,196 | 0,177 | 0,158 |
| $\mu_1(\alpha_1)$ | 249,90 | 5,47 | 5,44 | 8,05 | 7,65 |
| $\sigma_1(\beta_1)$ | 89,91 | 0,363 | 0,337 | 0,032 | 1702,23 |
| $\mu_2(\alpha_2)$ | 2147,76 | 7,58 | 7,33 | 2,19 | 4,38 |
| $\sigma_2(\beta_2)$ | 1639,40 | 0,558 | 0,739 | 0,001 | 7833,82 |
| stat. KS | 0,74 | 0,52 | 0,50 | 0,67 | 0,58 |
| p -KS | 0,665 | 1 | 1 | 0,820 | 1 |
| $\mathbb{E}_1(X)$ | 250 | 254 | 244 | 252 | 256 |
| $\sqrt{\text{Var}_1(X)}$ | 90 | 95 | 85 | 89 | 108 |
| $\mathbb{E}_2(X)$ | 2148 | 2289 | 2004 | 2190 | 2318 |
| $\sqrt{\text{Var}_2(X)}$ | 1639 | 1383 | 1708 | 1480 | 1502 |

la divisant par y^2 . Aussi nous pouvons voir que résoudre cette équation:

$$\frac{\partial}{\partial \theta} \log f_Y(y | \alpha, \beta) = \frac{\partial}{\partial \theta} \log \frac{1}{y^2} f_X(y^{-1} | \alpha, \beta) = 0$$

pour $\theta = (\alpha, \beta)$ est équivalent à résoudre cette équation:

$$\frac{\partial}{\partial \theta} \log f_X(y^{-1} | \alpha, \beta) = 0,$$

où

$$Y \sim IG(\alpha, \beta) \text{ et } X \sim G(\alpha, \beta).$$

Nous pouvons donc estimer les paramètres d'un mélange de lois inverse-gamma de la même façon que nous le faisons pour un mélange de lois gamma, en remplaçant simplement les données originales par leur inverse.

2.2.4. Exemple d'application

Nous avons appliqué l'algorithme EM pour ajuster les mélanges de lois vus à la section précédente aux données de l'étude ainsi que la méthode des moments pour le mélange de lois log-normales. Les résultats pour les premier et deuxième traitements apparaissent respectivement dans les tableaux 2.1 et 2.2. Les cinq premières rangées contiennent les estimations des paramètres de chaque modèle. Les deux rangées suivantes contiennent la statistique du test de Kolmogorov-Smirnov ainsi que la valeur- p correspondante. Le test de Kolmogorov-Smirnov

TABLEAU 2.2. Ajustement de distributions bimodales pour le traitement à la ciprofloxacine

| | normale | log-normale (EM) | log-normale (moments) | gamma | inverse-gamma |
|--------------------------|---------|---------------------|--------------------------|-------|---------------|
| ϵ | 0,155 | 0,139 | 0,204 | 0,152 | 0,094 |
| $\mu_1(\alpha_1)$ | 264,90 | 5,57 | 5,57 | 38,60 | 30,81 |
| $\sigma_1(\beta_1)$ | 39,98 | 0,169 | 0,161 | 0,145 | 8063,91 |
| $\mu_2(\alpha_2)$ | 439,07 | 6,05 | 5,89 | 9,21 | 13,51 |
| $\sigma_2(\beta_2)$ | 145,61 | 0,323 | 0,377 | 0,021 | 6244,11 |
| stat. KS | 1,18 | 1,31 | 1,33 | 1,26 | 1,34 |
| p -KS | 0,122 | 0,066 | 0,058 | 0,082 | 0,056 |
| $\mathbb{E}_1(Y)$ | 265 | 266 | 266 | 266 | 271 |
| $\sqrt{\text{Var}_1(Y)}$ | 40 | 45 | 43 | 43 | 50 |
| $\mathbb{E}_2(Y)$ | 439 | 447 | 388 | 439 | 499 |
| $\sqrt{\text{Var}_2(Y)}$ | 146 | 148 | 152 | 145 | 147 |

consiste à tester si les données proviennent de la distribution ajustée. Plus petite est la statistique et plus grande est la valeur- p , meilleur sera l'ajustement du modèle.

Nous voyons que l'ajustement bimodal de la loi log-normale génère les plus petites statistiques de Kolmogorov-Smirnov pour le premier traitement, soit respectivement 0,50 et 0,52 pour les méthodes des moments et l'algorithme EM, ce qui génère des valeurs- p de 1.

L'ajustement bimodal de la loi normale génère la plus petite statistique de Kolmogorov-Smirnov pour le deuxième traitement, soit 1,18 pour une valeur- p de 0,122. Les autres ajustements pour le deuxième traitement génèrent des valeurs- p entre 0,056 et 0,082.

Pour le reste du mémoire, nous choisirons l'ajustement par les lois log-normale bimodales, puisque les résultats sont adéquats et l'algorithme est plus simple à appliquer que pour une loi gamma ou inverse-gamma. Aussi il sera facile de déduire les développements pour un mélange de lois normales à partir d'un mélange de lois log-normales.

Finalement dans les quatre dernières rangées des deux tableaux, nous trouvons l'espérance et l'écart type des première et deuxième distributions de chaque mélange. Pour les calculer, nous nous servons des formules du tableau 2.3.

TABLEAU 2.3. Espérance et variance de distributions log-normale, gamma et inverse-gamma

| | $\mathbb{E}(X)$ | $\text{Var}(X)$ |
|----------------------------|--------------------------|---|
| $X \sim LN(\mu, \sigma^2)$ | $\exp[\mu + \sigma^2/2]$ | $\exp[2\mu + \sigma^2](\exp[\sigma^2] - 1)$ |
| $X \sim G(\alpha, \beta)$ | α/β | α/β^2 |
| $X \sim IG(\alpha, \beta)$ | $\beta/(\alpha - 1)$ | $\beta^2/((\alpha - 1)^2(\alpha - 2))$ |

2.3. MÉTHODE DE RÉÉCHANTILLONNAGE PARAMÉTRIQUE

Maintenant que nous avons ajusté un modèle paramétrique aux données de l'étude, nous pouvons effectuer le test bilatéral d'égalité des moyennes $H_0 : \mathbb{E}(X) = \mathbb{E}(Y)$ contre $H_1 : \mathbb{E}(X) \neq \mathbb{E}(Y)$ par la méthode de rééchantillonnage paramétrique, telle que décrite par Efron et Tibshirani (1993, section 16.5). Cette méthode est construite de la même façon que le test de Student et que la méthode de rééchantillonnage non paramétrique, soit en quatre étapes.

La première étape consiste à faire une hypothèse sur les distributions inconnues des variables aléatoires X et Y . Au lieu de supposer que X et Y ont une distribution normale comme le fait le test de Student, ou que les données proviennent de leur distribution empirique comme le fait la méthode de rééchantillonnage non paramétrique, nous supposons que X et Y suivent une loi ajustée aux données. Comme nous venons de le voir dans la section précédente, l'ajustement d'un mélange de lois log-normales aux données de l'étude est approprié. Les paramètres sont estimés par l'algorithme EM et nous utilisons les résultats des tableaux 2.1 et 2.2 pour poser:

$$\begin{aligned} f_X(x) &= (1 - \epsilon_x)f_1(x|\mu_{x_1}, \sigma_{x_1}^2) + \epsilon_x f_2(x|\mu_{x_2}, \sigma_{x_2}^2) \\ &= 0,838f_1(x | 5,47; 0,363^2) + 0,162f_2(x | 7,58; 0,558^2), \end{aligned}$$

$$\begin{aligned} f_Y(y) &= (1 - \epsilon_y)f_1(y|\mu_{y_1}, \sigma_{y_1}^2) + \epsilon_y f_2(y|\mu_{y_2}, \sigma_{y_2}^2) \\ &= 0,861f_1(y | 5,57; 0,169^2) + 0,139f_2(y | 6,05; 0,323^2), \end{aligned}$$

où $f_1(\cdot)$ et $f_2(\cdot)$ sont les fonctions de densité d'une log-normale.

Cette hypothèse est plus appropriée que celle du test de Student où nous supposons une loi normale. Elle présente aussi l'avantage de pouvoir utiliser une distribution continue, ce qui est plus réaliste que la distribution empirique supposée par la méthode de rééchantillonnage non paramétrique. Aussi, elle tient compte de la probabilité d'observer une valeur plus grande que la donnée maximale observée.

La deuxième étape est la même que celle du test de Student et de la méthode de rééchantillonnage non paramétrique, soit choisir la statistique

$$Z = \frac{(\bar{x} - \bar{y})}{\sqrt{(s_x^2/m) + (s_y^2/n)}},$$

et calculer $z_{obs}=1,92$.

La troisième étape consiste à trouver la distribution de la statistique Z sous l'hypothèse $H_0 : \mathbb{E}(X) = \mathbb{E}(Y)$. Comme pour le test de Student, il aurait pu être possible de trouver une distribution paramétrique de Z sous H_0 puisque nous avons supposé des distributions paramétriques pour X et Y . Mais il est impossible de le faire dû à la complexité de la distribution par rapport à la statistique. Nous approximerons donc la distribution de Z sous H_0 par simulation, un peu comme nous le faisons dans la méthode de rééchantillonnage non paramétrique. Chaque simulation b ($b = 1, \dots, B$) consiste à générer aléatoirement deux échantillons, notés $(x_1^{*b}, \dots, x_m^{*b})$ et $(y_1^{*b}, \dots, y_n^{*b})$, à partir des distributions supposées à la première étape, soient des mélanges de lois log-normales, et ce sous H_0 .

Pour y parvenir, nous générons une valeur à partir d'une loi $Bin(1, \epsilon)$. Si nous obtenons 0, nous générons une observation d'une loi $N(\mu_{x_1}, \sigma_{x_1}^2)$ et si nous obtenons 1, nous générons une observation d'une loi $N(\mu_{x_2}, \sigma_{x_2}^2)$. Puisque $X \sim LN(\mu, \sigma^2) \Leftrightarrow \log X \sim N(\mu, \sigma^2)$, nous prenons l'exponentielle de ce nombre. Nous avons donc généré une observation selon une loi $(1 - \epsilon_x)LN(\mu_{x_1}, \sigma_{x_1}^2) + \epsilon_x LN(\mu_{x_2}, \sigma_{x_2}^2)$. Nous répétons cela m fois pour générer $(x_1^{*b}, \dots, x_m^{*b})$. Nous refaisons le même processus pour le deuxième traitement à partir d'un mélange de lois log-normales $(1 - \epsilon_y)LN(\mu_{y_1}, \sigma_{y_1}^2) + \epsilon_y LN(\mu_{y_2}, \sigma_{y_2}^2)$ afin de générer $(y_1^{*b}, \dots, y_n^{*b})$.

Pour satisfaire l'hypothèse $H_0 : \mathbb{E}(X) = \mathbb{E}(Y)$, nous translatons l'échantillon de $(\mathbb{E}(X) - \mathbb{E}(Y))$ afin d'obtenir $(y_1^{**b} + \mathbb{E}(X) - \mathbb{E}(Y), \dots, y_n^{**b} + \mathbb{E}(X) - \mathbb{E}(Y)) = (y_1^{*b}, \dots, y_n^{*b})$, où

$$\mathbb{E}(X) = (1 - \epsilon_x)e^{\mu_{x_1} + \sigma_{x_1}^2/2} + \epsilon_x e^{\mu_{x_2} + \sigma_{x_2}^2/2}$$

et

$$\mathbb{E}(Y) = (1 - \epsilon_y)e^{\mu_{y_1} + \sigma_{y_1}^2/2} + \epsilon_y e^{\mu_{y_2} + \sigma_{y_2}^2/2}.$$

Nous calculons ensuite la réplique de la statistique Z , soit

$$Z^{*b} = \frac{(\bar{x}^{*b} - \bar{y}^{*b})}{\sqrt{(s_x^{2*b}/m) + (s_y^{2*b}/n)}},$$

qui est calculée à partir des deux nouveaux échantillons. Cette simulation est faite indépendamment pour $b = 1, \dots, B$ afin d'obtenir Z^{*1}, \dots, Z^{*B} .

La quatrième étape est identique à celle de la méthode de rééchantillonnage non paramétrique, soit l'approximation de la valeur- p pour le test bilatéral par la formule suivante:

$$\hat{p} = \# \{|Z^*| \geq |z_{obs}|\} / B.$$

Dans l'étude, pour $B = 10\,000$ simulations, 1 090 valeurs de Z^* sont plus petites que $-|z_{obs}| = -1,92$ tandis que 50 valeurs de Z^* sont plus grandes que $|z_{obs}| = 1,92$. Alors

$$\hat{p} = (1\,090 + 50) / 10\,000 = 0,114.$$

Nous nous rappellerons que le \hat{p} du rééchantillonnage non paramétrique était de 0,155 tandis que le \hat{p} du test de Student était de 0,061.

Si nous voulons effectuer le test unilatéral $H_0 : \mathbb{E}(X) = \mathbb{E}(Y)$ contre $H_1 : \mathbb{E}(X) > \mathbb{E}(Y)$, alors

$$\begin{aligned} \hat{p} &= \begin{cases} \# \{Z^* \geq z_{obs}\} / B & ; \text{ si } z_{obs} > 0, \\ \# \{Z^* \leq z_{obs}\} / B & ; \text{ si } z_{obs} < 0, \end{cases} \\ &= 0,005. \end{aligned}$$

Nous nous rappellerons que pour le test unilatéral, le \hat{p} du rééchantillonnage non paramétrique était de 0,0034 tandis que le \hat{p} du test de Student était de 0,0305.

2.4. MÉTHODE DE RÉÉCHANTILLONNAGE MODIFIÉ

La méthode de rééchantillonnage modifié est une version modifiée de la méthode de rééchantillonnage non paramétrique, où les données sont supposées provenir de la distribution empirique, c'est-à-dire la distribution de X est définie par $\mathbb{P}(x) = 1/m$ si $x \in \{x_1, \dots, x_m\}$. Dans la méthode de rééchantillonnage modifié, nous voulons tenir compte de la probabilité, notée λ_x , d'observer une valeur plus grande que $\max\{x_1, \dots, x_m\}$, c'est-à-dire:

$$\lambda_x = \mathbb{P}[X \geq \max\{x_1, \dots, x_m\}] = \mathbb{P}[X > x_{(m)}] = 1 - F(x_{(m)}),$$

où $x_{(m)} = \max\{x_1, \dots, x_m\}$ et $F(\cdot)$ est la fonction de répartition paramétrique supposée.

Ainsi au lieu de générer aléatoirement avec remise des échantillons à partir des données observées, nous générons à partir de la distribution suivante:

$$\mathbb{P}(X \in A) = \begin{cases} (1 - \lambda_x)/m & ; \text{ si } A = x, x \in \{x_1, \dots, x_m\}, \\ \lambda_x & ; \text{ si } A = (x_{(m)}, \infty). \end{cases}$$

Nous arrivons à générer des données à partir d'une telle distribution par l'algorithme suivant:

- 1) nous générons une valeur u_1 selon une loi uniforme $U(0, \frac{m}{1-\lambda_x})$,
- 2) nous définissons $u_2 = [u_1] + 1$, où $[u_1]$ est la partie entière de u_1 ,
- 3) si $1 \leq u_2 \leq m$, la valeur générée est x_{u_2} ,
- 4) si $u_2 \geq m + 1$, nous générons une valeur x plus grande que $x_{(m)}$.

Il reste à déterminer comment générer une valeur x plus grande que $x_{(m)}$. C'est la partie paramétrique de la méthode de rééchantillonnage modifié. C'est à partir d'une distribution ajustée aux données, dont la fonction de densité est donnée par $f(x)$, que nous générons une donnée plus grande que $x_{(m)}$. En général, nous avons

$$f(x | X > x_{(m)}) = \begin{cases} 0 & ; \text{ si } x \leq x_{(m)}, \\ \frac{f(x)}{1-F(x_{(m)})} & ; \text{ si } x > x_{(m)}, \end{cases}$$

et donc

$$F(x | X > x_{(m)}) = \begin{cases} 0 & ; \text{ si } x \leq x_{(m)}, \\ \frac{F(x) - F(x_{(m)})}{1 - F(x_{(m)})} & ; \text{ si } x > x_{(m)}. \end{cases}$$

Afin de générer une valeur de x selon cette distribution, nous générons u_3 selon une loi uniforme $U(0, 1)$. Puisque $F(x | X > x_{(m)}) \sim U(0, 1)$, nous avons donc

$$\begin{aligned} u_3 &= \frac{F(x) - F(x_{(m)})}{1 - F(x_{(m)})} \\ \Rightarrow F(x) &= F(x_{(m)}) + u_3(1 - F(x_{(m)})) \\ &= F(x_{(m)})(1 - u_3) + u_3 \\ \Rightarrow x &= F^{-1}[F(x_{(m)})(1 - u_3) + u_3], \end{aligned}$$

où x est une valeur plus grande que $x_{(m)}$. Si nous choisissons d'ajuster une loi bimodale, nous avons le modèle suivant:

$$f(x) = (1 - \epsilon_x) f_1(x) + \epsilon_x f_2(x)$$

et donc

$$F(x) = (1 - \epsilon_x) F_1(x) + \epsilon_x F_2(x).$$

Si $x > x_{(m)}$ et si les deux modes sont suffisamment éloignés, alors nous pouvons supposer que $F_1(x) \approx 1$. Ainsi si nous supposons que $F_1(x) = 1$ si $x > x_{(m)}$, alors

$$F(x) = (1 - \epsilon_x) + \epsilon_x F_2(x), \quad x > x_{(m)}.$$

Nous avons alors

$$\lambda_x = 1 - F(x_{(m)}) = \epsilon_x(1 - F_2(x_{(m)}))$$

et l'équation

$$F(x) = F(x_{(m)})(1 - u_3) + u_3$$

devient

$$\begin{aligned}
 (1 - \epsilon_x) + \epsilon_x F_2(x) &= [(1 - \epsilon_x) + \epsilon_x F_2(x_{(m)})](1 - u_3) + u_3 \\
 \Leftrightarrow F_2(x) &= F_2(x_{(m)}) + u_3 - u_3 F_2(x_{(m)}) \\
 \Leftrightarrow x &= F_2^{-1}[F_2(x_{(m)})(1 - u_3) + u_3].
 \end{aligned}$$

Il ne reste plus qu'à déterminer $F_2(x)$ et à calculer $F_2(x_{(m)})$. À la section 2.2, nous avons choisi d'ajuster des mélanges de lois log-normales aux données des deux traitements. Pour le premier traitement, nous avons le modèle suivant:

$$\begin{aligned}
 f_X(x|\epsilon_x, \mu_{x_1}, \sigma_{x_1}^2, \mu_{x_2}, \sigma_{x_2}^2) &= (1 - \epsilon_x)f_1(x|\mu_{x_1}, \sigma_{x_1}^2) + \epsilon_x f_2(x|\mu_{x_2}, \sigma_{x_2}^2) \\
 &= (1 - 0,162)f_1(x | 5,47; 0,363^2) \\
 &\quad + 0,162f_2(x | 7,58; 0,558^2)
 \end{aligned}$$

et pour le deuxième traitement, nous avons le modèle suivant:

$$\begin{aligned}
 f_Y(y|\epsilon_y, \mu_{y_1}, \sigma_{y_1}^2, \mu_{y_2}, \sigma_{y_2}^2) &= (1 - \epsilon_y)f_1(y|\mu_{y_1}, \sigma_{y_1}^2) + \epsilon_y f_2(y|\mu_{y_2}, \sigma_{y_2}^2) \\
 &= (1 - 0,139)f_1(y | 5,57; 0,169^2) \\
 &\quad + 0,139f_2(y | 6,05; 0,323^2),
 \end{aligned}$$

où $f_1(\cdot)$ et $f_2(\cdot)$ sont les fonctions de densité d'une loi log-normale. Nous avons observé

$$x_{(m)} = \max(x_1, \dots, x_m) = 6182 \text{ pour le premier traitement}$$

et

$$y_{(n)} = \max(y_1, \dots, y_n) = 743 \text{ pour le deuxième traitement.}$$

Nous pouvons vérifier que $F_1(x_{(m)}) \approx F_1(y_{(n)}) \approx 1$. Effectivement, $1 - F_1(x_{(m)}) < 1,0\text{E-}16$ et $1 - F_1(y_{(n)}) = 3,77\text{E-}10$. Aussi

$$\begin{aligned} F_2(x_{(m)}) &= \mathbb{P}(LN(7,58; 0,558^2) \leq 6182) \\ &= \mathbb{P}(N(7,58; 0,558^2) \leq \log 6182) \\ &= 0,9803 \end{aligned}$$

et

$$\begin{aligned} F_2(y_{(n)}) &= \mathbb{P}(LN(6,05; 0,323^2) \leq 743) \\ &= \mathbb{P}(N(6,05; 0,323^2) \leq \log 743) \\ &= 0,9587. \end{aligned}$$

Nous pouvons aussi calculer

$$\begin{aligned} \lambda_x &= \epsilon_x(1 - F_2(x_{(m)})) = 0,0032, \\ \lambda_y &= \epsilon_y(1 - F_2(y_{(n)})) = 0,0057. \end{aligned}$$

Nous générons donc $x, x > x_{(m)}$, par cette relation:

$$x = \exp[7,58 + 0,558 \Phi^{-1}(0,9803(1 - u_3) + u_3)] \quad (2.4.1)$$

où u_3 est générée selon une loi uniforme $U(0, 1)$ et $\Phi^{-1}(\cdot)$ est le quantile d'une $N(0, 1)$. De la même façon, nous générons $y, y > y_{(n)}$, par cette relation:

$$y = \exp[6,05 + 0,323 \Phi^{-1}(0,9587(1 - u_3) + u_3)].$$

Nous sommes donc en mesure de générer les deux échantillons $(x_1^{*b}, \dots, x_m^{*b})$ et $(y_1^{**b}, \dots, y_n^{**b})$. Le reste de l'algorithme est identique à celui de la méthode de rééchantillonnage paramétrique, sauf pour le calcul de $\mathbb{E}(X) - \mathbb{E}(Y)$ dans la translation de l'échantillon $(y_1^{**b}, \dots, y_n^{**b})$ dans le but d'obtenir $(y_1^{*b}, \dots, y_n^{*b})$.

En effet, dans le cas de la méthode de rééchantillonnage modifié, nous avons

$$\begin{aligned}\mathbb{E}(X) &= \bar{x}(1 - \lambda_x) + \lambda_x \frac{\int_{x_{(m)}}^{\infty} x f(x) dx}{1 - F(x_{(m)})} \\ &= \bar{x}(1 - \lambda_x) + \lambda_x \mathbb{E}(X | X > x_{(m)})\end{aligned}$$

où $\mathbb{E}(X | X > x_{(m)})$ peut être évaluée par la méthode de simulations Monte Carlo en générant x_i selon l'équation (2.4.1) un grand nombre de fois, disons S , afin de générer la suite x_1, \dots, x_S . Nous estimons ensuite l'intégrale de cette façon:

$$\mathbb{E}(X | X > x_{(m)}) \approx \frac{1}{S} \sum_{i=1}^S x_i.$$

Le même procédé est utilisé pour calculer $\mathbb{E}(Y)$.

Nous avons appliqué cette méthode aux données de l'étude et après 10 000 simulations, nous trouvons $p=0,162$ pour le test bilatéral et $p=0,0034$ pour le test unilatéral.

Dans ce chapitre, nous avons vu comment estimer les paramètres d'un modèle de mélange de deux lois, par la méthode des moments et par l'algorithme EM. Cette dernière s'avère une méthode efficace et relativement simple à appliquer et elle est par conséquent choisie pour ajuster les mélanges de lois dans ce mémoire.

Nous avons appliqué l'algorithme EM pour estimer les paramètres de mélange de lois normales, log-normales, gamma et inverse-gamma. Les résultats d'ajustement, jugés par le test de Kolmogorov-Smirnov, s'avèrent satisfaisant pour le mélange de distributions log-normales, surtout pour le premier traitement. De plus, l'algorithme est plus facile à appliquer que pour les mélanges de distributions gamma ou inverse-gamma, et par conséquent le mélange de distributions log-normales ajusté par l'algorithme EM est choisi pour modéliser les données de l'étude.

Finalement, nous avons présenté deux autres méthodes de rééchantillonnage, l'une appelée rééchantillonnage paramétrique et l'autre rééchantillonnage modifié. La première est semblable à la méthode de rééchantillonnage non paramétrique, à la différence qu'à la première étape, nous supposons que les distributions des

TABLEAU 2.4. Résultats des méthodes de rééchantillonnage

| | non-paramétrique | paramétrique | modifié |
|-------------|------------------|--------------|---------|
| valeur- p | 0,155 | 0,114 | 0,162 |

coûts sont celles que nous avons préalablement ajustées aux données. Aussi à la troisième étape, nous générons selon ces distributions.

La deuxième est la méthode de rééchantillonnage modifié, qui est aussi semblable à la méthode de rééchantillonnage non paramétrique, à la différence que la probabilité d'observer une valeur supérieure à la plus grande donnée observée est considérée.

Ces deux méthodes ont l'avantage de supposer des distributions encore plus réalistes que la méthode de rééchantillonnage non paramétrique, mais elles ont le désavantage de gagner en complexité. Le tableau 2.4 rappelle les résultats des trois méthodes de rééchantillonnage. Nous voyons peu de différence entre les valeurs- p de ces méthodes.

CHAPITRE 3

APPROCHE BAYÉSIENNE

Dans ce chapitre, nous présentons l'approche bayésienne pour comparer deux traitements concurrents. À partir du mélange de deux lois log-normales pour modéliser la distribution des coûts de chacun des traitements tel que décrit au chapitre 2, l'approche bayésienne requiert le choix de lois *a priori* pour les paramètres de ces distributions. Au lieu de considérer ces paramètres fixes comme nous le faisons dans la théorie fréquentiste, nous les considérons stochastiques. Nous sommes ensuite intéressés à trouver la distribution *a posteriori* de ces paramètres, qui combine l'information des lois *a priori* et des observations.

À partir de cette distribution *a posteriori* des paramètres, nous pouvons répondre au test d'hypothèses unilatéral ou bilatéral d'égalité des moyennes des chapitres 1 et 2 de façon bayésienne. Pour ce faire, nous calculerons la probabilité *a posteriori* que l'hypothèse H_0 survienne. Une autre avenue possible est de calculer la probabilité *a posteriori* que le premier traitement soit moins coûteux que le deuxième traitement.

À la section 3.1, nous présentons les tests d'hypothèses selon l'approche bayésienne. Nous y voyons la fonction des coûts, la fonction de perte *a posteriori* et la règle de décision qui minimise cette dernière fonction.

À la section 3.2, nous introduisons les familles de lois conjuguées et les familles exponentielle de lois.

À la section 3.3, nous voyons une méthode d'estimation des paramètres que nous appelons méthode bayésienne. C'est une méthode qui utilise l'algorithme EM afin de trouver les hyperparamètres qui maximisent la fonction de densité *a posteriori*. Cette méthode est appliquée à l'étude et la méthode de rééchantillonnage paramétrique est effectuée à partir de ces estimés.

Aux sections 3.4 et 3.5, nous présentons les tests d'hypothèses unilatéral et bilatéral selon l'approche bayésienne. Dans ces sections, seuls les paramètres de position sont considérés comme étant des variables aléatoires.

Aux sections 3.6 et 3.7, nous présentons aussi les tests d'hypothèses unilatéral et bilatéral selon l'approche bayésienne, mais en considérant tous les paramètres comme étant des variables aléatoires, incluant les paramètres de dispersion et de poids.

À la section 3.8, nous appliquons les méthodes décrites aux sections précédentes sur les données de l'étude considérée et nous y présentons les résultats.

Finalement à la section 3.9, nous présentons une méthode pour calculer la probabilité *a posteriori* que le premier traitement soit moins coûteux que le deuxième traitement.

3.1. TESTS D'HYPOTHÈSES BAYÉSIENS

Soit le modèle statistique $f(x|\theta), \theta \in \Theta$. Nous voulons tester $H_0 : \theta \in \Theta_0$ contre $H_1 : \theta \in \Theta_1$, où Θ_0 et Θ_1 sont des sous-ensembles d'intérêt de Θ tels que $\Theta_0 \cup \Theta_1 = \Theta$ et $\Theta_0 \cap \Theta_1 = \emptyset$. Ce test revient à estimer la fonction indicatrice

$$I_{\Theta_0}(\theta) = \begin{cases} 1 & ; \text{ si } \theta \in \Theta_0, \\ 0 & ; \text{ si } \theta \notin \Theta_0. \end{cases}$$

Si $\Theta_1 \neq \Theta_0^c$, nous supposons que le complémentaire de $\Theta_0 \cup \Theta_1$ n'a pas d'intérêt pour le problème, c'est-à-dire $\mathbb{P}^\pi[\theta \in (\Theta_0 \cup \Theta_1)^c] = 0$. Pour estimer $I_{\Theta_0}(\theta)$, il suffit de disposer d'une fonction de coûts adéquate. Nous envisagerons une fonction de coûts qui pénalise différemment les mauvaises décisions selon que l'hypothèse

nulle est vraie ou fausse, définie par:

$$L(\theta, \gamma) = \begin{cases} a_0 & ; \text{ si } \theta \in \Theta_0 \text{ et } \gamma = 0, \\ a_1 & ; \text{ si } \theta \notin \Theta_0 \text{ et } \gamma = 1, \\ 0 & ; \text{ sinon,} \end{cases}$$

où γ est un estimateur de $I_{\Theta_0}(\theta)$, c'est-à-dire

$$\gamma = \begin{cases} 0 & ; \text{ si nous décidons de rejeter } H_0, \\ 1 & ; \text{ sinon.} \end{cases}$$

DÉFINITION 3.1.1. *Si nous connaissons la loi a posteriori de θ , soit $\pi(\theta|x)$, alors nous pouvons trouver l'estimateur γ , noté $\gamma^\pi(x)$, qui minimise la fonction de perte a posteriori, définie par*

$$\rho(\pi, \gamma|x) = \mathbb{E}^{\pi(\theta|x)}[L(\theta, \gamma)] = \int_{\Theta} L(\theta, \gamma)\pi(\theta|x) d\theta.$$

Nous avons donc deux valeurs possibles pour la fonction de perte a posteriori, soient:

$$\rho(\pi, \gamma|x) = \begin{cases} a_0 \mathbb{P}^\pi(\theta \in \Theta_0|x) & ; \text{ si } \gamma = 0, \\ a_1 \mathbb{P}^\pi(\theta \notin \Theta_0|x) & ; \text{ si } \gamma = 1. \end{cases}$$

Si $\mathbb{P}^\pi(\theta \in \Theta_0|x) > \frac{a_1}{a_0+a_1}$, alors $\rho(\pi, 1|x) < \rho(\pi, 0|x)$ et nous choisissons $\gamma = 1$ puisque c'est la décision qui minimise la fonction de perte a posteriori. Alors nous définissons

$$\gamma^\pi(x) = \begin{cases} 1 & ; \text{ si } \mathbb{P}^\pi(\theta \in \Theta_0|x) > \frac{a_1}{a_0+a_1}, \\ 0 & ; \text{ sinon.} \end{cases}$$

Nous sommes donc intéressés à calculer $\mathbb{P}^\pi(\theta \in \Theta_0|x)$, la probabilité a posteriori de l'hypothèse nulle $H_0 : \theta \in \Theta_0$. Cette valeur est un peu l'analogue de la valeur- p du test classique et $\frac{a_1}{a_0+a_1}$ l'analogue du niveau α . Par exemple, si nous choisissons $a_0 = 19$ et $a_1 = 1$, c'est-à-dire que nous pénalisons beaucoup plus l'erreur de type I, soit rejeter H_0 quand H_0 est vraie, nous rejeterons H_0 si $\mathbb{P}^\pi(\theta \in \Theta_0|x) < 0,05$.

Dans ce mémoire, nous sommes intéressés à tester $H_0 : \mathbb{E}(X) = \mathbb{E}(Y)$ contre $H_1 : \mathbb{E}(X) \neq \mathbb{E}(Y)$ pour le test bilatéral d'égalité des moyennes et à tester $H_0 : \mathbb{E}(X) = \mathbb{E}(Y)$ contre $H_1 : \mathbb{E}(X) > \mathbb{E}(Y)$ pour le test unilatéral. Nous aborderons les deux tests distinctement puisque l'approche est quelque peu différente.

Rappelons que le modèle de mélange de deux lois a été défini au chapitre 2 de cette façon:

$$f_X(x|\epsilon_x, \boldsymbol{\theta}_{x_1}, \boldsymbol{\theta}_{x_2}) = (1 - \epsilon_x)f_1(x|\boldsymbol{\theta}_{x_1}) + \epsilon_x f_2(x|\boldsymbol{\theta}_{x_2})$$

et

$$f_Y(y|\epsilon_y, \boldsymbol{\theta}_{y_1}, \boldsymbol{\theta}_{y_2}) = (1 - \epsilon_y)f_1(y|\boldsymbol{\theta}_{y_1}) + \epsilon_y f_2(y|\boldsymbol{\theta}_{y_2}).$$

Puisque nous avons choisi de modéliser les données par un mélange de lois log-normales, le modèle peut s'écrire ainsi:

$$f_X(x|\epsilon_x, \mu_{x_1}, \sigma_{x_1}^2, \mu_{x_2}, \sigma_{x_2}^2) = (1 - \epsilon_x)f_1(x | \mu_{x_1}, \sigma_{x_1}^2) + \epsilon_x f_2(x | \mu_{x_2}, \sigma_{x_2}^2)$$

et

$$f_Y(y|\epsilon_y, \mu_{y_1}, \sigma_{y_1}^2, \mu_{y_2}, \sigma_{y_2}^2) = (1 - \epsilon_y)f_1(y | \mu_{y_1}, \sigma_{y_1}^2) + \epsilon_y f_2(y | \mu_{y_2}, \sigma_{y_2}^2),$$

où $f_1(\cdot)$ et $f_2(\cdot)$ sont les fonctions de densité d'une distribution log-normale.

Dans cette section, nous considérons $\epsilon_x, \sigma_{x_1}^2, \sigma_{x_2}^2, \epsilon_y, \sigma_{y_1}^2, \sigma_{y_2}^2$ comme étant connues. Si ce n'est pas le cas, nous les estimons par les valeurs trouvées à l'aide de l'algorithme EM décrit au chapitre 2, soient

$$(\hat{\epsilon}_x, \hat{\sigma}_{x_1}^2, \hat{\sigma}_{x_2}^2, \hat{\epsilon}_y, \hat{\sigma}_{y_1}^2, \hat{\sigma}_{y_2}^2) = (0,162; 0,363^2; 0,558^2; 0,139; 0,169^2; 0,323^2).$$

Ainsi les seuls paramètres considérés comme des variables aléatoires sont $\mu_{x_1}, \mu_{x_2}, \mu_{y_1}$ et μ_{y_2} . Nous voyons dans la prochaine section comment déterminer les lois *a priori* sur des paramètres.

3.2. LOIS CONJUGUÉES

DÉFINITION 3.2.1. *Une famille Γ de lois sur Θ est dite conjuguée pour $f(x|\theta)$*

$$\text{si } \forall \pi(\theta) \in \Gamma, \pi(\theta|x) \in \Gamma.$$

Par exemple, si $X \sim \text{Bin}(n, \theta)$, alors la famille de lois $\Gamma = \{\text{Be}(\alpha, \beta) | \alpha > 0, \beta > 0\}$ est dite conjuguée pour $f(x|\theta)$ car

$$\begin{aligned} \theta &\sim \text{Be}(\alpha, \beta) \\ \Rightarrow \pi(\theta|x) &\propto f(x|\theta)\pi(\theta) \\ &\propto [\theta^x(1-\theta)^{n-x}] [\theta^{\alpha-1}(1-\theta)^{\beta-1}] \\ &= \theta^{\alpha+x-1}(1-\theta)^{\beta+n-x-1} \\ \Rightarrow \theta|x &\sim \text{Be}(\alpha+x, \beta+n-x) \in \Gamma. \end{aligned}$$

Si $f(x|\theta)$ fait partie d'une famille exponentielle de lois, alors il existe un théorème permettant de trouver la famille de lois conjuguées correspondantes. Voyons la définition de famille exponentielle.

DÉFINITION 3.2.2. *Soit*

$$C : \Theta \rightarrow \mathbb{R}_+, H : \mathcal{X} \rightarrow \mathbb{R}_+, T : \mathcal{X} \rightarrow \mathbb{R}^k, R : \Theta \rightarrow \mathbb{R}^k$$

et $A \subset \mathbb{R}^k$, A indépendant de θ , alors la famille de densités

$$\mathcal{F} = \{f(x|\theta) | \theta \in \Theta\}$$

est dite une famille exponentielle de lois si

$$f(x|\theta) = C(\theta)H(x) \exp \left\{ \sum_{i=1}^k R_i(\theta)T_i(x) \right\} \mathbb{I}_A(x)$$

où

$$\mathbb{I}_A(x) = \begin{cases} 1 & ; \text{ si } x \in A, \\ 0 & ; \text{ sinon.} \end{cases}$$

Voyons maintenant comment trouver la famille de lois conjuguées correspondant à une loi appartenant à une famille exponentielle (voir Bickel et Doksum (1977, section 2.3)).

THÉORÈME 3.2.1. Si $f(x|\theta)$ appartient à une famille exponentielle de lois, alors une famille conjuguée par $f(x|\theta)$ est donnée par

$$\Gamma = \pi(\theta|\boldsymbol{\lambda}) = K(\boldsymbol{\lambda})C^{\lambda_0}(\theta) \exp \left\{ \sum_{i=1}^k \lambda_i R_i(\theta) \right\}$$

où $\boldsymbol{\lambda} = (\lambda_0, \lambda_1, \dots, \lambda_k)^t$ est tel que $\int_{\Theta} \pi(\theta|\boldsymbol{\lambda}) d\theta = 1$ et $K(\boldsymbol{\lambda})$ est une constante de normalisation qui ne dépend pas de θ , $\theta \in \Theta$, $\lambda \in \mathbb{R}^{k+1}$.

Appliquons maintenant ce théorème à l'étude. Nous avons déterminé pour le premier traitement que pour chacune des deux distributions du mélange, $x|\mu, \sigma^2 \sim LN(x|\mu, \sigma^2)$ pour $(\mu, \sigma^2) = (\mu_{x_i}, \sigma_{x_i}^2)$, $i = 1, 2$, $-\infty < \mu < \infty$, $\sigma > 0$. La même chose s'applique pour le deuxième traitement en changeant x par y . Montrons d'abord que $f(x|\mu, \sigma^2)$ fait partie d'une famille exponentielle et trouvons ensuite la famille de lois pour μ, σ^2 conjuguées par $f(x|\mu, \sigma^2)$.

$$\begin{aligned} f(x|\mu, \sigma^2) &= \frac{1}{x\sqrt{2\pi\sigma^2}} \exp \left[-\frac{1}{2\sigma^2} (\log x - \mu)^2 \right] \\ &= \frac{1}{x\sqrt{2\pi\sigma^2}} \exp \left[-\frac{(\log x)^2}{2\sigma^2} + \frac{\mu \log x}{\sigma^2} - \frac{\mu^2}{2\sigma^2} \right] \\ &= \left(\frac{1}{\sqrt{\sigma^2}} \exp \left[-\frac{\mu^2}{2\sigma^2} \right] \right) \left(\frac{1}{x\sqrt{2\pi}} \right) \\ &\times \exp \left[\left(-\frac{1}{\sigma^2} \right) \left(\frac{(\log x)^2}{2} \right) + \left(\frac{\mu}{\sigma^2} \right) (\log x) \right] \\ &= C(\mu, \sigma^2) H(x) \exp \left\{ \sum_{i=1}^2 R_i(\mu, \sigma^2) T_i(x) \right\} I_A(x), \end{aligned}$$

où

$$\begin{aligned}
C(\mu, \sigma^2) &= \left(\frac{1}{\sqrt{\sigma^2}} \exp \left[-\frac{\mu}{2\sigma^2} \right] \right) \in \mathbb{R}_+, \\
H(x) &= \left(\frac{1}{x\sqrt{2\pi}} \right) \in \mathbb{R}_+, \\
R_1(\mu, \sigma^2) &= \left(-\frac{1}{\sigma^2} \right), \\
R_2(\mu, \sigma^2) &= \left(\frac{\mu}{\sigma^2} \right), \\
T_1(x) &= \left(\frac{(\log x)^2}{2} \right), \\
T_2(x) &= (\log x), \\
A = \mathbb{R}_+ &\text{ est indépendant de } \mu \text{ et } \sigma^2.
\end{aligned}$$

Alors par le théorème 3.2.1, puisque $f(x|\mu, \sigma^2)$ est une fonction de densité d'une loi log-normale et donc appartient à une famille exponentielle de lois, alors une famille conjuguée par $f(x|\mu, \sigma^2)$ est donnée par

$$\Gamma = \pi(\mu, \sigma^2 | \boldsymbol{\lambda}) = K(\boldsymbol{\lambda}) C^{\lambda_0}(\mu, \sigma^2) \exp \left\{ \sum_{i=1}^k \lambda_i R_i(\mu, \sigma^2) \right\},$$

c'est-à-dire

$$\begin{aligned}
\pi(\mu, \sigma^2 | \boldsymbol{\lambda}) &= K(\boldsymbol{\lambda}) \left(\frac{1}{\sqrt{\sigma^2}} \exp \left[-\frac{\mu^2}{2\sigma^2} \right] \right)^{\lambda_0} \exp \left[\lambda_1 \left(-\frac{1}{\sigma^2} \right) + \lambda_2 \left(\frac{\mu}{\sigma^2} \right) \right] \\
&= K(\boldsymbol{\lambda}) \frac{1}{(\sqrt{\sigma^2})^{\lambda_0}} \exp \left[-\frac{\lambda_1}{\sigma^2} \right] \exp \left[-\frac{\lambda_0 \mu^2}{2\sigma^2} + \frac{\lambda_2 \mu}{\sigma^2} \right] \\
&= K(\boldsymbol{\lambda}) \frac{1}{\sqrt{\sigma^2}} \exp \left[-\frac{\lambda_0}{2\sigma^2} \left(\mu - \frac{\lambda_2}{\lambda_0} \right)^2 \right] \\
&\times \exp \left[\frac{\lambda_2^2}{2\lambda_0 \sigma^2} \right] \frac{1}{(\sqrt{\sigma^2})^{\lambda_0 - 1}} \exp \left[-\frac{\lambda_1}{\sigma^2} \right] \\
&= K(\boldsymbol{\lambda}) \left\{ \frac{1}{\sqrt{\sigma^2}} \exp \left[-\frac{\lambda_0}{2\sigma^2} \left(\mu - \frac{\lambda_2}{\lambda_0} \right)^2 \right] \right\} \\
&\times \left\{ \frac{1}{(\sigma^2)^{(\lambda_0 - 1)/2}} \exp \left[-\frac{1}{\sigma^2} \left(\lambda_1 - \frac{\lambda_2^2}{2\lambda_0} \right) \right] \right\}.
\end{aligned}$$

Ainsi, nous trouvons que

$$\mu | \sigma^2 \sim N \left(\frac{\lambda_2}{\lambda_0}, \frac{\sigma^2}{\lambda_0} \right)$$

et

$$\sigma^2 \sim IG\left(\frac{\lambda_0 - 3}{2}, \lambda_1 - \frac{\lambda_2^2}{2\lambda_0}\right),$$

ou de façon équivalente

$$\mu \mid \sigma^2 \sim N(\mu_0, c_0\sigma^2)$$

et

$$\sigma^2 \sim IG(\alpha, \beta).$$

3.3. ESTIMATION DES PARAMÈTRES PAR LA MÉTHODE BAYÉSIENNE

La méthode bayésienne consiste à trouver les paramètres qui maximisent le logarithme de la fonction de densité *a posteriori* des paramètres, soit

$$\log \pi(\boldsymbol{\psi} \mid \mathbf{y}) = \log \left[\frac{f(\mathbf{y} \mid \boldsymbol{\psi})\pi(\boldsymbol{\psi})}{m(\mathbf{y})} \right] = \log [f(\mathbf{y} \mid \boldsymbol{\psi})\pi(\boldsymbol{\psi})] - \log m(\mathbf{y}),$$

où $f(\mathbf{y} \mid \boldsymbol{\psi})$ est la fonction de vraisemblance, $\pi(\boldsymbol{\psi})$ la fonction de densité *a priori* des paramètres $\boldsymbol{\psi}$ et $m(\mathbf{y})$ la fonction de densité marginale de \mathbf{y} .

Cette méthode se réduit à la méthode du maximum de vraisemblance si $\pi(\boldsymbol{\psi}) = 1$, c'est-à-dire si les lois *a priori* sur $\boldsymbol{\psi}$ sont non informatives. Toutefois, le modèle appliqué à l'étude est un mélange de lois comportant des variables latentes notées \mathbf{z} , avec la fonction de densité définie par:

$$f(y_i \mid z_i, \boldsymbol{\psi}) = f(y_i \mid z_i, \boldsymbol{\theta}_1, \boldsymbol{\theta}_2, \epsilon) = f_1(y_i \mid \boldsymbol{\theta}_1)^{1-z_i} f_2(y_i \mid \boldsymbol{\theta}_2)^{z_i}$$

où $z_i \mid \epsilon \sim \text{Bin}(1, \epsilon)$, $i = 1, \dots, n$. Comme les données complètes sont données par \mathbf{y}, \mathbf{z} , nous chercherons $\boldsymbol{\psi}$ qui maximise:

$$\begin{aligned} \log \pi(\boldsymbol{\psi} \mid \mathbf{y}, \mathbf{z}) &= \log \left(\frac{f(\mathbf{y}, \mathbf{z} \mid \boldsymbol{\psi})\pi(\boldsymbol{\psi})}{m(\mathbf{y}, \mathbf{z})} \right) = \log [f(\mathbf{y} \mid \mathbf{z}, \boldsymbol{\psi})\pi_{\mathbf{z}}(\mathbf{z} \mid \boldsymbol{\psi})\pi(\boldsymbol{\psi})] \\ &\quad - \log m(\mathbf{y}, \mathbf{z}), \end{aligned}$$

où $\boldsymbol{\psi} = (\boldsymbol{\theta}_1, \boldsymbol{\theta}_2, \epsilon)^t$. Nous utiliserons l'algorithme EM vu au chapitre 2 pour résoudre ce problème. L'algorithme EM consiste à générer une séquence d'estimateurs $\{\boldsymbol{\psi}^{(m)}\}$ à partir d'un estimateur initial $\boldsymbol{\psi}^{(0)}$ jusqu'à convergence vers $\hat{\boldsymbol{\psi}}$. Chaque itération notée m consiste en deux étapes:

1) Étape E (Espérance)

Calculer $Q(\boldsymbol{\psi}, \boldsymbol{\psi}^{(m)}) = \mathbb{E}^{\mathbf{z} | \boldsymbol{\psi}^{(m)}, \mathbf{y}} [\log[f(\mathbf{y} | \mathbf{z}, \boldsymbol{\psi})\pi(\mathbf{z}, \boldsymbol{\psi})\pi(\boldsymbol{\psi})]]$.

2) Étape M (Maximisation)

Trouver $\boldsymbol{\psi} = \boldsymbol{\psi}^{(m+1)}$ qui maximise $Q(\boldsymbol{\psi}, \boldsymbol{\psi}^{(m)})$ en résolvant l'équation

$$\frac{\partial}{\partial \boldsymbol{\psi}} Q(\boldsymbol{\psi}, \boldsymbol{\psi}^{(m)}) = 0.$$

Appliquons maintenant cet algorithme à l'ajustement d'un mélange de deux lois log-normales. Rappelons que l'ajustement d'un mélange de deux lois log-normales se fait de la même façon en utilisant le logarithme des données au lieu des données originales. Tout d'abord, nous avons vu à la section 2.2.2 que pour un mélange de lois en général nous avons les deux résultats suivants:

$$z_i | \epsilon, \boldsymbol{\theta}_1, \boldsymbol{\theta}_2, \mathbf{y} \sim B \left(1, \frac{\epsilon f_2(y_i | \boldsymbol{\theta}_2)}{(1 - \epsilon) f_1(y_i | \boldsymbol{\theta}_1) + \epsilon f_2(y_i | \boldsymbol{\theta}_2)} \right)$$

et

$$\begin{aligned} Q(\boldsymbol{\psi}, \boldsymbol{\psi}^{(m)}) &= \sum_{i=1}^n [1 - \mathbb{E}^{z_i | \boldsymbol{\psi}^{(m)}, y_i}(z_i)] \log f_1(y_i | \boldsymbol{\theta}_1) \\ &+ \sum_{i=1}^n \mathbb{E}^{z_i | \boldsymbol{\psi}^{(m)}, y_i}(z_i) \log f_2(y_i | \boldsymbol{\theta}_2) + \log(1 - \epsilon) \sum_{i=1}^n [1 - \mathbb{E}^{z_i | \boldsymbol{\psi}^{(m)}, y_i}(z_i)] \\ &+ \log \epsilon \sum_{i=1}^n \mathbb{E}^{z_i | \boldsymbol{\psi}^{(m)}, y_i}(z_i) + \log \pi(\epsilon) + \log \pi(\boldsymbol{\theta}_1) + \log \pi(\boldsymbol{\theta}_2). \end{aligned}$$

Pour le cas particulier du mélange de lois log-normales, nous avons pour $\boldsymbol{\theta}_j = (\mu_j, \sigma_j^2)$, $j = 1, 2$,

$$\begin{aligned} \log f_j(y_i | \boldsymbol{\theta}_j) &= \log f_j(y_i | \mu_j, \sigma_j^2) \\ &= \log \frac{1}{y_i \sqrt{2\pi\sigma_j^2}} \exp \left[-\frac{1}{2\sigma_j^2} (\log y_i - \mu_j)^2 \right] \\ &= -\log y_i - \frac{1}{2} \log(2\pi) - \frac{1}{2} \log \sigma_j^2 - \frac{1}{2\sigma_j^2} (\log y_i - \mu_j)^2. \end{aligned}$$

Il faut finalement déterminer les lois *a priori* pour $\boldsymbol{\psi} = \epsilon, \mu_1, \sigma_1^2, \mu_2, \sigma_2^2$ avant d'appliquer l'algorithme EM. Nous les choisirons de façon à ce qu'elles soient conjuguées, par rapport à $\mathbf{z} | \epsilon \sim \text{Bin}(1, \epsilon)$ pour ϵ et par rapport à $y_i | \mu_j, \sigma_j^2 \sim \text{LN}(\mu_j, \sigma_j^2)$ pour μ_j, σ_j^2 , $j = 1, 2$. Ainsi les lois *a priori* et leurs fonctions de

densité sont:

$$\begin{aligned}\epsilon &\sim Be(c, d) \text{ et } \pi(\epsilon) = \frac{\Gamma(c+d)}{\Gamma(c)\Gamma(d)} \epsilon^{c-1} (1-\epsilon)^{d-1}, \\ &0 < \epsilon < 1, c, d > 0,\end{aligned}$$

$$\begin{aligned}\mu_j &\sim N(\lambda_j, \tau_j^2) \text{ et } \pi(\mu_j) = \frac{1}{\sqrt{2\pi\tau_j^2}} \exp\left[-\frac{1}{2\tau_j^2}(\mu_j - \lambda_j)^2\right], \\ &-\infty < \mu_j, \lambda_j < \infty, \tau_j^2 > 0, j = 1, 2,\end{aligned}$$

$$\begin{aligned}\sigma_j^2 &\sim IG(\alpha_j, \beta_j) \text{ et } \pi(\sigma_j^2) = \frac{\beta_j^{\alpha_j}}{\Gamma(\alpha_j)} \frac{1}{(\sigma_j^2)^{\alpha_j+1}} \exp\left[-\frac{\beta_j}{\sigma_j^2}\right], \\ &\sigma_j^2 > 0, \beta_j > 0, \alpha_j > 2, j = 1, 2.\end{aligned}$$

Effectivement, nous avons vu à la section 3.2 que la loi *a priori* bêta pour ϵ est conjuguée par rapport à la loi binômiale de paramètre ϵ , et que les lois normales pour μ et inverse-gamma pour σ^2 sont conjuguées par rapport à la loi log-normale de paramètre μ et σ^2 .

À partir d'estimés préliminaires $\boldsymbol{\psi}^{(m)} = (\epsilon^{(m)}, \mu_1^{(m)}, \sigma_1^{2(m)}, \mu_2^{(m)}, \sigma_2^{2(m)})$, nous sommes maintenant en mesure d'appliquer l'algorithme EM. Nous trouvons donc:

$$\begin{aligned}\mathbb{E}^{z_i | \boldsymbol{\psi}^{(m)}, y_i}(z_i) &= \mathbb{E}^{z_i | \boldsymbol{\psi}^{(m)}, y_i}(z_i) \\ &= \frac{\epsilon^{(m)} f_2(y_i | \mu_2^{(m)}, \sigma_2^{2(m)})}{(1 - \epsilon^{(m)}) f_1(y_i | \mu_1^{(m)}, \sigma_1^{2(m)}) + \epsilon^{(m)} f_2(y_i | \mu_2^{(m)}, \sigma_2^{2(m)})},\end{aligned}$$

où $f_1(\cdot)$ et $f_2(\cdot)$ sont les fonctions de densité d'une loi log-normale.

Nous calculons $\epsilon^{(m+1)}$ en résolvant l'équation suivante:

$$\begin{aligned}0 &= \frac{\partial}{\partial \epsilon} Q(\boldsymbol{\psi}, \boldsymbol{\psi}^{(m)}) \\ &= \frac{\partial}{\partial \epsilon} \left[\log(1 - \epsilon) \sum_{i=1}^n [1 - \mathbb{E}^{z_i | \boldsymbol{\psi}^{(m)}, y_i}(z_i)] + \log \epsilon \sum_{i=1}^n \mathbb{E}^{z_i | \boldsymbol{\psi}^{(m)}, y_i}(z_i) \right. \\ &\quad \left. + (c - 1) \log \epsilon + (d - 1) \log(1 - \epsilon) \right] \\ &= \frac{-\sum_{i=1}^n (1 - \mathbb{E}^{z_i | \boldsymbol{\psi}^{(m)}, y_i}(z_i)) - (d - 1)}{1 - \epsilon} + \frac{\sum_{i=1}^n \mathbb{E}^{z_i | \boldsymbol{\psi}^{(m)}, y_i}(z_i) + (c - 1)}{\epsilon}.\end{aligned}$$

Nous obtenons alors la proposition suivante.

PROPOSITION 3.3.1.

$$\epsilon^{(m+1)} = \frac{\sum_{i=1}^n \mathbb{E}^{z_i | \psi^{(m)}, y_i}(z_i) + c - 1}{n + c + d - 2}.$$

Nous remarquons que si la loi *a priori* de ϵ est une $Be(1, 1) = U(0, 1)$, c'est-à-dire $c = d = 1$, alors $\epsilon^{(m+1)} = \frac{\sum_{i=1}^n \mathbb{E}^{z_i | \psi^{(m)}, y_i}(z_i)}{n}$ comme dans la méthode du maximum de vraisemblance (MMV) par l'algorithme EM. Tout le poids est alors consacré aux données. À l'autre extrême, si $c, d \rightarrow \infty$, alors $\text{Var}(\epsilon) = \frac{cd}{(c+d+1)(c+d)^2} \rightarrow 0$ et $\epsilon^{(m+1)} \rightarrow \frac{c}{c+d} = \mathbb{E}(\epsilon)$, c'est-à-dire tout le poids est donné à l'information *a priori* sur ϵ .

Ensuite, nous calculons $\mu_2^{(m+1)}$ en résolvant l'équation suivante:

$$\begin{aligned} 0 &= \frac{\partial}{\partial \mu_2} Q(\psi, \psi^{(m)}) \\ &= \frac{\partial}{\partial \mu_2} \left[- \sum_{i=1}^n \mathbb{E}^{z_i | \psi^{(m)}, y_i}(z_i) \left(\frac{1}{2\sigma_2^2} (y_i - \mu_2)^2 \right) - \frac{1}{2\tau_2^2} (\mu_2 - \lambda_2)^2 \right] \\ &= \frac{\sum_{i=1}^n \mathbb{E}^{z_i | \psi^{(m)}, y_i}(z_i) (y_i - \mu_2)}{\sigma_2^2} - \frac{(\mu_2 - \lambda_2)}{\tau_2^2}. \end{aligned}$$

Nous obtenons alors la proposition suivante.

PROPOSITION 3.3.2.

$$\mu_2^{(m+1)} = \frac{\tau_2^2 \sum_{i=1}^n \mathbb{E}^{z_i | \psi^{(m)}, y_i}(z_i) y_i + \sigma_b^{2(m)} \lambda_2}{\tau_2^2 \sum_{i=1}^n \mathbb{E}^{z_i | \psi^{(m)}, y_i}(z_i) + \sigma_2^{2(m)}}.$$

Nous remarquons que si la loi *a priori* de μ_2 est une $N(\lambda_2, \tau_2^2)$ et $\tau_2^2 \rightarrow \infty$, alors $\mu_2^{(m+1)} \rightarrow \frac{\sum_{i=1}^n \mathbb{E}^{z_i | \psi^{(m)}, y_i}(z_i) y_i}{\sum_{i=1}^n \mathbb{E}^{z_i | \psi^{(m)}, y_i}(z_i)}$ comme dans la méthode du maximum de vraisemblance par l'algorithme EM et tout le poids est consacré aux données. Aussi si $\tau_2^2 \rightarrow 0$, alors $\mu_2^{(m+1)} \rightarrow \lambda_2 = \mathbb{E}(\mu_2)$, c'est-à-dire tout le poids est consacré à l'information *a priori*. Nous trouvons ensuite $\sigma_2^{2(m+1)}$ en résolvant l'équation

suivante:

$$\begin{aligned}
0 &= \frac{\partial}{\partial \sigma_2^2} Q(\psi, \psi^{(m)}) \\
&= \frac{\partial}{\partial \sigma_2^2} \left[\sum_{i=1}^n \mathbb{E}^{z_i | \psi^{(m)}, y_i}(z_i) \left(-\frac{1}{2} \log \sigma_2^2 - \frac{1}{2\sigma_2^2} (y_i - \mu_2)^2 \right) \right. \\
&\quad \left. - (\alpha_2 + 1) \log \sigma_2^2 - \frac{\beta_2}{\sigma_2^2} \right] \\
&= \sum_{i=1}^n \mathbb{E}^{z_i | \psi^{(m)}, y_i}(z_i) \left(-\frac{1}{2\sigma_2^2} + \frac{(y_i - \mu_2)^2}{2(\sigma_2^2)^2} \right) - \frac{\alpha_2 + 1}{\sigma_2^2} + \frac{\beta_2}{(\sigma_2^2)^2}.
\end{aligned}$$

Nous obtenons alors la proposition suivante.

PROPOSITION 3.3.3.

$$\sigma_2^{2(m+1)} = \frac{\frac{1}{2} \sum_{i=1}^n \mathbb{E}^{z_i | \psi^{(m)}, y_i}(z_i) (y_i - \mu_2^{(m+1)})^2 + \beta_2}{\frac{1}{2} \sum_{i=1}^n \mathbb{E}^{z_i | \psi^{(m)}, y_i}(z_i) + \alpha_2 + 1}.$$

De la même façon, nous trouvons les deux propositions suivantes.

PROPOSITION 3.3.4.

$$\mu_1^{(m+1)} = \frac{\tau_1^2 \sum_{i=1}^n (1 - \mathbb{E}^{z_i | \psi^{(m)}, y_i}(z_i)) y_i + \sigma_1^{2(m)} \lambda_1}{\tau_1^2 \sum_{i=1}^n (1 - \mathbb{E}^{z_i | \psi^{(m)}, y_i}(z_i)) + \sigma_1^{2(m)}}.$$

PROPOSITION 3.3.5.

$$\sigma_1^{2(m+1)} = \frac{\frac{1}{2} \sum_{i=1}^n (1 - \mathbb{E}^{z_i | \psi^{(m)}, y_i}(z_i)) (y_i - \mu_1^{(m+1)})^2 + \beta_1}{\frac{1}{2} \sum_{i=1}^n (1 - \mathbb{E}^{z_i | \psi^{(m)}, y_i}(z_i)) + \alpha_1 + 1}.$$

Nous remarquons que si $\beta_2 = k(\alpha_2 - 1)$ où $k > 0$ est fixe, si la loi *a priori* de σ_2^2 est une $IG(\alpha_2, \beta_2)$ et si $\alpha_2 \rightarrow \infty$ (et donc $\beta_2 = k(\alpha_2 - 1) \rightarrow \infty$), alors

$$\text{Var}(\sigma_2^2) = \frac{\beta_2^2}{(\alpha_2 - 1)^2(\alpha_2 - 2)} = \frac{k^2(\alpha_2 - 1)^2}{(\alpha_2 - 1)^2(\alpha_2 - 2)} = \frac{k^2}{\alpha_2 - 2} \rightarrow 0$$

et

$$\sigma_2^{2(m+1)} \rightarrow k = \frac{\beta_2}{\alpha_2 - 1} = \mathbb{E}(\sigma_2^2)$$

et tout le poids est consacré à l'information *a priori*. À l'autre extrême, si $\beta_2 = k(\alpha_2 - 1)$ où $k > 0$ fixe et $\alpha_2 \rightarrow 2^+$, alors $\text{var}(\sigma_2^2) = \frac{k^2}{\alpha_2 - 2} \rightarrow \infty$ et $\sigma_2^{2(m+1)} \rightarrow \frac{\sum_{i=1}^n \mathbb{E}^{z_i | \psi^{(m)}, y_i}(z_i) (y_i - \mu_2^{(m+1)})^2 + 2k}{\sum_{i=1}^n \mathbb{E}^{z_i | \psi^{(m)}, y_i}(z_i) + 6}$, ce qui donne très peu de poids à l'information *a priori*.

Nous avons vu que la méthode bayésienne d'estimation des paramètres requiert le choix de lois *a priori* et de leurs paramètres. Pour la loi normale, il est naturel de fixer les paramètres μ et σ^2 puisqu'ils représentent la moyenne et la variance. Pour les lois bêta et inverse-gamma, il est plus difficile de fixer les paramètres directement. Nous pouvons toutefois choisir les moyennes et variances de ces distributions puis en déduire les paramètres. Si $U \sim Be(c, d)$, nous avons:

$$\begin{aligned}\mathbb{E}(U) &= \mu = \frac{c}{c+d}, \\ \text{Var}(U) &= \sigma^2 = \frac{cd}{(c+d+1)(c+d)^2},\end{aligned}$$

et donc nous pouvons en déduire que:

$$\begin{aligned}c &= \frac{(1-\mu)\mu^2 - \mu\sigma^2}{\sigma^2}, \\ d &= c \left(\frac{1-\mu}{\mu} \right),\end{aligned}$$

et si $U \sim IG(\alpha, \beta)$, nous avons:

$$\begin{aligned}\mathbb{E}(U) &= \mu = \frac{\beta}{\alpha-1}, \\ \text{Var}(U) &= \sigma^2 = \frac{\beta^2}{(\alpha-1)^2(\alpha-2)},\end{aligned}$$

et donc nous pouvons en déduire que:

$$\begin{aligned}\alpha &= \frac{\mu^2}{\sigma^2} + 2, \\ \beta &= \mu \left(\frac{\mu^2}{\sigma^2} + 1 \right).\end{aligned}$$

Nous avons appliqué cet algorithme aux données de l'étude en choisissant des lois *a priori* identiques pour les deux traitements puisque nous les considérons équivalents au point de départ. Les espérances et écart types des paramètres des lois *a priori* sont affichés au tableau 3.1 et les résultats de l'ajustement sont affichés au tableau 3.2.

À la section 2.3, nous avons introduit la méthode de rééchantillonnage paramétrique. Cette méthode requiert l'ajustement de distributions et nous pouvons donc nous servir des estimés du tableau 3.2 pour modéliser les données. Nous avons appliqué la méthode de rééchantillonnage paramétrique aux données de l'étude et nous avons obtenu $z_{obs} = 1,92$. Nous avons effectué 10 000 simulations (B), et nous avons observé que 1 216 valeurs de Z^* sont plus petites que

TABLEAU 3.1. Choix des paramètres des lois *a priori* pour l'ajustement d'un mélange de lois log-normales par la méthode bayésienne

| | $\mathbb{E}(\epsilon)$ | $\sqrt{\text{Var}(\epsilon)}$ | $\mathbb{E}(\mu_i)$ | $\sqrt{\text{Var}(\mu_i)}$ | $\mathbb{E}(\sigma_i^2)$ | $\sqrt{\text{Var}(\sigma_i^2)}$ |
|---------|------------------------|-------------------------------|---------------------|----------------------------|--------------------------|---------------------------------|
| | 0,20 | 0,05 | | | | |
| $i = 1$ | | | 5,5 | 0,25 | 0,25 | 0,1 |
| $i = 2$ | | | 6,5 | 0,25 | 0,5 | 0,2 |

TABLEAU 3.2. Résultats de l'ajustement d'un mélange de lois log-normales par la méthode bayésienne

| | ϵ | μ_1 | σ_1 | μ_2 | σ_2 |
|--------------|------------|---------|------------|---------|------------|
| traitement 1 | 0,195 | 5,45 | 0,385 | 6,89 | 0,736 |
| traitement 2 | 0,156 | 5,61 | 0,303 | 6,00 | 0,592 |

$-|z_{obs}| = -1,92$ tandis que 44 valeurs de Z^* sont plus grandes que $|z_{obs}| = 1,92$.

Alors

$$\hat{p} = (1\ 216 + 44)/10\ 000 = 0,126.$$

Nous nous rappellerons que le \hat{p} du rééchantillonnage paramétrique avec les estimés trouvés précédemment était de 0,114. La valeur- p pour le test unilatéral est de 0,0044.

3.4. TEST D'HYPOTHÈSES UNILATÉRAL AVEC POIDS ET VARIANCES CONNUS

Dans cette section, les seuls paramètres considérés comme étant des variables aléatoires sont $\mu_{x_1}, \mu_{x_2}, \mu_{y_1}, \mu_{y_2}, \mathbf{z}$ et \mathbf{w} , les paramètres de poids (ϵ_x, ϵ_y) et de dispersion ($\sigma_{x_1}^2, \sigma_{x_2}^2, \sigma_{y_1}^2, \sigma_{y_2}^2$) sont considérés connus. Afin d'alléger le texte, il arrivera souvent que nous mentionnerons seulement les formules pour le traitement 1. Les formules pour le traitement 2 seront alors identiques en changeant \mathbf{x} par \mathbf{y} et \mathbf{z} par \mathbf{w} .

Le modèle pour le traitement 1 a été défini de cette façon:

$$f_X(x_i|\mu_{x_1}, \mu_{x_2}) = (1 - \epsilon_x)f_{X,1}(x_i|\mu_{x_1}) + \epsilon_x f_{X,2}(x_i|\mu_{x_2}), \quad i = 1, \dots, m,$$

où

$$f_{X,j}(x_i|\mu_{x_j}) = \frac{1}{x_i \sqrt{2\pi\sigma_{x_j}^2}} \exp \left[-\frac{1}{2\sigma_{x_j}^2} (\log x_i - \mu_{x_j})^2 \right], \quad j = 1, 2,$$

soient les fonctions de densité d'une distribution log-normale de paramètres μ_{x_1} , $\sigma_{x_1}^2$ et de paramètres μ_{x_2} , $\sigma_{x_2}^2$.

Nous pouvons récrire le modèle comme suit:

$$f_X(x_i | \mu_{x_1}, \mu_{x_2}, z_i) = f_{X,1}(x_i | \mu_{x_1})^{(1-z_i)} f_{X,2}(x_i | \mu_{x_2})^{z_i}, \quad i = 1, \dots, m,$$

où

$$z_i \sim \text{Bin}(1, \epsilon_x),$$

c'est-à-dire

$$f_X(x_i | \mu_{x_1}, \mu_{x_2}, z_i) = \begin{cases} f_{X,1}(x_i | \mu_{x_1}) & ; \text{ avec probabilité } 1 - \epsilon_x, \\ f_{X,2}(x_i | \mu_{x_2}) & ; \text{ avec probabilité } \epsilon_x. \end{cases}$$

La fonction de probabilité de z_i est

$$\pi_Z(z_i) = (1 - \epsilon_x)^{(1-z_i)} (\epsilon_x)^{z_i},$$

et donc pour $\mathbf{z} = (z_1, \dots, z_m)^t$,

$$\pi_{\mathbf{Z}}(\mathbf{z}) = \prod_{i=1}^m \pi_Z(z_i) = g_{\mathbf{Z},1}(\mathbf{z}) g_{\mathbf{Z},2}(\mathbf{z}), \quad (3.4.1)$$

où

$$g_{\mathbf{Z},1}(\mathbf{z}) = (1 - \epsilon_x)^{\sum (1-z_i)}, \quad (3.4.2)$$

$$g_{\mathbf{Z},2}(\mathbf{z}) = (\epsilon_x)^{\sum z_i}. \quad (3.4.3)$$

Cette notation nous sera utile un peu plus loin dans le chapitre. Aussi pour $\mathbf{x} = (x_1, \dots, x_m)^t$, nous avons:

$$\begin{aligned} f_{\mathbf{X}}(\mathbf{x} | \mu_{x_1}, \mu_{x_2}, \mathbf{z}) &= \prod_{i=1}^m f_X(x_i | \mu_{x_1}, \mu_{x_2}, z_i) \\ &= \prod_{i=1}^m \left(\frac{1}{x_i \sqrt{2\pi} \sigma_{x_1}} \exp \left[-\frac{1}{2\sigma_{x_1}^2} (\log x_i - \mu_{x_1})^2 \right] \right)^{(1-z_i)} \\ &\quad \times \left(\frac{1}{x_i \sqrt{2\pi} \sigma_{x_2}} \exp \left[-\frac{1}{2\sigma_{x_2}^2} (\log x_i - \mu_{x_2})^2 \right] \right)^{z_i} \\ &= K_{\mathbf{X},1}(\mathbf{x}) g_{\mathbf{X},1}(\mathbf{x} | \mu_{x_1}, \mathbf{z}) g_{\mathbf{X},2}(\mathbf{x} | \mu_{x_2}, \mathbf{z}), \end{aligned} \quad (3.4.4)$$

où

$$g_{\mathbf{X},1}(\mathbf{x} \mid \mu_{x_1}, \mathbf{z}) = \frac{1}{(\sigma_{x_1})^{(m-\sum_i z_i)}} \exp \left[-\frac{1}{2\sigma_{x_1}^2} \sum_{i=1}^m (\log x_i - \mu_{x_1})^2 (1 - z_i) \right], \quad (3.4.5)$$

$$g_{\mathbf{X},2}(\mathbf{x} \mid \mu_{x_2}, \mathbf{z}) = \frac{1}{(\sigma_{x_2})^{\sum_i z_i}} \exp \left[-\frac{1}{2\sigma_{x_2}^2} \sum_{i=1}^m (\log x_i - \mu_{x_2})^2 z_i \right], \quad (3.4.6)$$

$$K_{\mathbf{X},1}(\mathbf{x}) = \frac{1}{(2\pi)^{m/2} \prod_{i=1}^m x_i}. \quad (3.4.7)$$

Il faut maintenant choisir une loi *a priori* pour μ_{x_1} et μ_{x_2} . Nous avons vu à la section 3.2 que la famille de lois normales pour les paramètres μ_{x_1} et μ_{x_2} est conjuguée pour la famille de lois log-normales. Ainsi nous choisissons $\mu_{x_1} \sim N(\lambda_{x_1}, \tau_{x_1}^2)$ et $\mu_{x_2} \sim N(\lambda_{x_2}, \tau_{x_2}^2)$. Aussi, dans le mélange de distributions log-normales, nous voulons que la moyenne de la première distribution soit plus petite que la deuxième, c'est-à-dire $\mathbb{E}_1(X) < \mathbb{E}_2(X)$, où

$$\mathbb{E}_j(X) = e^{\mu_{x_j} + \sigma_{x_j}^2/2}, \quad j = 1, 2.$$

Nous définissons alors la fonction de densité conjointe de (μ_{x_1}, μ_{x_2}) comme suit:

$$\pi_{\mu_{\mathbf{x}}}(\mu_{x_1}, \mu_{x_2}) = K_{\mathbf{X},2}(\mathbf{x}) \pi_{\mu_{x_1}}(\mu_{x_1}) \pi_{\mu_{x_2}}(\mu_{x_2}) \mathbb{I}_X(\mu_{x_1}, \mu_{x_2}) \quad (3.4.8)$$

où

$$K_{\mathbf{X},2}(\mathbf{x})^{-1} = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \pi_{\mu_{x_1}}(\mu_{x_1}) \pi_{\mu_{x_2}}(\mu_{x_2}) \mathbb{I}_X(\mu_{x_1}, \mu_{x_2}) d\mu_{x_1} d\mu_{x_2}, \quad (3.4.9)$$

$$\pi_{\mu_{x_j}}(\mu_{x_j}) = \frac{1}{\sqrt{2\pi}\tau_{x_j}} \exp \left[-\frac{1}{2\tau_{x_j}^2} (\mu_{x_j} - \lambda_{x_j})^2 \right], \quad j = 1, 2, \quad (3.4.10)$$

$$\mathbb{I}_X(\mu_{x_1}, \mu_{x_2}) = \begin{cases} 1 & ; \text{ si } \mathbb{E}_1(X) < \mathbb{E}_2(X), \\ 0 & ; \text{ sinon.} \end{cases} \quad (3.4.11)$$

Afin d'alléger la notation, nous écrivons $\mu_{\mathbf{x}}$ pour (μ_{x_1}, μ_{x_2}) . Rappelons que nous avons les mêmes résultats pour le traitement 2 en changeant \mathbf{x} par \mathbf{y} et \mathbf{z} par \mathbf{w} .

Le test d'hypothèses unilatéral est le suivant:

$$H_0 : \mathbb{E}(X) < \mathbb{E}(Y) \text{ contre } H_1 : \mathbb{E}(X) \geq \mathbb{E}(Y),$$

où

$$\mathbb{E}(X) = (1 - \epsilon_x)e^{\mu_{x_1} + \sigma_{x_1}^2/2} + \epsilon_x e^{\mu_{x_2} + \sigma_{x_2}^2/2}, \quad (3.4.12)$$

$$\mathbb{E}(Y) = (1 - \epsilon_y)e^{\mu_{y_1} + \sigma_{y_1}^2/2} + \epsilon_y e^{\mu_{y_2} + \sigma_{y_2}^2/2}. \quad (3.4.13)$$

Comme nous avons observé $\mathbf{x} = (x_1, \dots, x_m)^t$ et $\mathbf{y} = (y_1, \dots, y_n)^t$, nous sommes intéressés à calculer, tel que vu à la section 3.1, la probabilité *a posteriori* que l'hypothèse H_0 soit vraie, c'est-à-dire:

$$\begin{aligned} p(\mathbf{x}, \mathbf{y}) &= \mathbb{P}[H_0 : \mathbb{E}(X) < \mathbb{E}(Y) \mid \mathbf{x}, \mathbf{y}] \\ &= \mathbb{E}^{\pi(\mu_{\mathbf{x}}, \mu_{\mathbf{y}}, \mathbf{z}, \mathbf{w} \mid \mathbf{x}, \mathbf{y})} (I_{XY}(\mu_{\mathbf{x}}, \mu_{\mathbf{y}})) \\ &= \sum_{\mathbf{z}} \sum_{\mathbf{w}} \int_{\mu_{\mathbf{x}}} \int_{\mu_{\mathbf{y}}} I_{XY}(\mu_{\mathbf{x}}, \mu_{\mathbf{y}}) \pi(\mu_{\mathbf{x}}, \mu_{\mathbf{y}}, \mathbf{z}, \mathbf{w} \mid \mathbf{x}, \mathbf{y}) d\mu_{\mathbf{y}} d\mu_{\mathbf{x}} \end{aligned}$$

où

$$I_{XY}(\mu_{\mathbf{x}}, \mu_{\mathbf{y}}) = \begin{cases} 1 & ; \text{ si } \mathbb{E}(X) < \mathbb{E}(Y), \\ 0 & ; \text{ sinon,} \end{cases}$$

et

$$\begin{aligned} \pi(\mu_{\mathbf{x}}, \mu_{\mathbf{y}}, \mathbf{z}, \mathbf{w} \mid \mathbf{x}, \mathbf{y}) &= \\ &= \frac{f_{\mathbf{X}, \mathbf{Y}}(\mathbf{x}, \mathbf{y} \mid \mu_{\mathbf{x}}, \mu_{\mathbf{y}}, \mathbf{z}, \mathbf{w}) \pi(\mu_{\mathbf{x}}, \mu_{\mathbf{y}}, \mathbf{z}, \mathbf{w})}{\sum_{\mathbf{z}} \sum_{\mathbf{w}} \int_{\mu_{\mathbf{x}}} \int_{\mu_{\mathbf{y}}} f_{\mathbf{X}, \mathbf{Y}}(\mathbf{x}, \mathbf{y} \mid \mu_{\mathbf{x}}, \mu_{\mathbf{y}}, \mathbf{z}, \mathbf{w}) \pi(\mu_{\mathbf{x}}, \mu_{\mathbf{y}}, \mathbf{z}, \mathbf{w}) d\mu_{\mathbf{y}} d\mu_{\mathbf{x}}}. \end{aligned}$$

Afin d'alléger la notation, nous écrivons $\int_{\mu_{\mathbf{x}}}$ pour $\int_{\mu_{x_1}} \int_{\mu_{x_2}}$ et $d\mu_{\mathbf{x}}$ pour $d\mu_{x_1} d\mu_{x_2}$. La même notation s'applique pour le traitement 2 en changeant \mathbf{x} par \mathbf{y} .

Nous résoudrons cette intégrale par la méthode de simulation Monte Carlo avec fonctions d'importance. La méthode est décrite comme suit:

Soit l'intégrale suivante à résoudre:

$$\begin{aligned} p(\mathbf{x}, \mathbf{y}) &= \int_{\Theta} q(\theta) \pi(\theta \mid \mathbf{x}, \mathbf{y}) d\theta \\ &= \frac{\int_{\Theta} q(\theta) f(\mathbf{x}, \mathbf{y} \mid \theta) \pi(\theta) d\theta}{\int_{\Theta} f(\mathbf{x}, \mathbf{y} \mid \theta) \pi(\theta) d\theta} \\ &= \frac{\int_{\Theta} q(\theta) w(\theta) I(\theta) d\theta}{\int_{\Theta} w(\theta) I(\theta) d\theta}, \end{aligned}$$

où

$$w(\theta) = \frac{f(\mathbf{x}, \mathbf{y} \mid \theta) \pi(\theta)}{I(\theta)}.$$

Ainsi nous approximons $p(\mathbf{x}, \mathbf{y})$ par

$$\hat{p}(\mathbf{x}, \mathbf{y}) = \frac{\sum_{i=1}^S q(\boldsymbol{\theta}_i) w(\boldsymbol{\theta}_i)}{\sum_{i=1}^S w(\boldsymbol{\theta}_i)}$$

où $\boldsymbol{\theta}_i, i = 1, \dots, S$, sont générées indépendamment à partir de $I(\boldsymbol{\theta})$, qui représente la fonction d'importance à partir de laquelle nous sommes en mesure de générer facilement les $\boldsymbol{\theta}_i$. Il est démontré par Robert (1992, section 9.2) que $\hat{p}(\mathbf{x}, \mathbf{y})$ tend vers $p(\mathbf{x}, \mathbf{y})$ quand S tend vers l'infini.

Pour l'intégrale que nous voulons résoudre,

$\boldsymbol{\theta}$ correspond à $\mu_{\mathbf{x}}, \mu_{\mathbf{y}}, \mathbf{z}, \mathbf{w}$,

$q(\boldsymbol{\theta})$ correspond à $I_{XY}(\mu_{\mathbf{x}}, \mu_{\mathbf{y}})$,

$I(\boldsymbol{\theta})$ correspond à $I(\mu_{\mathbf{x}}, \mu_{\mathbf{y}}, \mathbf{z}, \mathbf{w}) = I_{\mu_{\mathbf{x}}}(\mu_{\mathbf{x}}) I_{\mu_{\mathbf{y}}}(\mu_{\mathbf{y}}) I_{\mathbf{Z}}(\mathbf{z}) I_{\mathbf{W}}(\mathbf{w})$.

Trouvons maintenant $I_{\mu_{\mathbf{x}}}(\mu_{\mathbf{x}})$ et $I_{\mathbf{Z}}(\mathbf{z})$ et les résultats seront les mêmes pour $I_{\mu_{\mathbf{y}}}(\mu_{\mathbf{y}})$ et $I_{\mathbf{W}}(\mathbf{w})$ en changeant \mathbf{x} par \mathbf{y} et \mathbf{z} par \mathbf{w} .

Comme le suggère Robert (1992, section 9.2), nous essaierons de poser $I_{\mu_{\mathbf{x}}}(\mu_{\mathbf{x}})$ et $I_{\mathbf{Z}}(\mathbf{z})$ le plus près possible des distributions *a posteriori* de $\mu_{\mathbf{x}}$ et \mathbf{z} , soit $\pi_{\mu_{\mathbf{x}}}(\mu_{\mathbf{x}} | \mathbf{x})$ et $\pi_{\mathbf{Z}}(\mathbf{z} | \mathbf{x})$ en développant $\pi_{\mu_{\mathbf{x}}}(\mu_{\mathbf{x}} | \mathbf{x}, \mathbf{z})$ et $\pi_{\mathbf{Z}}(\mathbf{z} | \mathbf{x}, \mu_{\mathbf{x}})$.

Tout d'abord,

$$\begin{aligned} \pi_{\mu_{\mathbf{x}}}(\mu_{\mathbf{x}} | \mathbf{x}, \mathbf{z}) &\propto f_{\mathbf{X}, \mathbf{Z}}(\mathbf{x}, \mathbf{z} | \mu_{\mathbf{x}}) \pi_{\mu_{\mathbf{x}}}(\mu_{\mathbf{x}}) \\ &\propto f_{\mathbf{X}}(\mathbf{x} | \mathbf{z}, \mu_{\mathbf{x}}) \pi_{\mu_{\mathbf{x}}}(\mu_{\mathbf{x}}) \\ &\propto \exp \left[-\frac{1}{2\sigma_{x_1}^2} \sum_{i=1}^m (\log x_i - \mu_{x_1})^2 (1 - z_i) \right] \exp \left[-\frac{1}{2\tau_{x_1}^2} (\mu_{x_1} - \lambda_{x_1})^2 \right] \\ &\times \exp \left[-\frac{1}{2\sigma_{x_2}^2} \sum_{i=1}^m (\log x_i - \mu_{x_2})^2 z_i \right] \exp \left[-\frac{1}{2\tau_{x_2}^2} (\mu_{x_2} - \lambda_{x_2})^2 \right] \\ &\times I_X(\mu_{x_1}, \mu_{x_2}). \end{aligned} \tag{3.4.14}$$

Nous développons cette expression à l'annexe 1, ce qui nous amène à poser

$$I_{\mu_{\mathbf{x}}}(\mu_{\mathbf{x}}) = \pi_{\mu_{x_1}}(\mu_{x_1} | \mathbf{x}, \mathbf{z}) \pi_{\mu_{x_2}}(\mu_{x_2} | \mathbf{x}, \mathbf{z}), \tag{3.4.15}$$

où

$$\begin{aligned} \pi_{\mu_{x_1}}(\mu_{x_1} | \mathbf{x}, \mathbf{z}) = & \frac{\sqrt{\tau_{x_1}^2 \sum (1 - z_i) + \sigma_{x_1}^2}}{\sqrt{2\pi\tau_{x_1}\sigma_{x_1}}} \\ & \times \exp \left[-\frac{\tau_{x_1}^2 \sum (1 - z_i) + \sigma_{x_1}^2}{2\tau_{x_1}^2 \sigma_{x_1}^2} \left(\mu_{x_1} - \frac{\bar{x}_{(1)}\tau_{x_1}^2 \sum (1 - z_i) + \lambda_{x_1}\sigma_{x_1}^2}{\tau_{x_1}^2 \sum (1 - z_i) + \sigma_{x_1}^2} \right)^2 \right], \end{aligned} \quad (3.4.16)$$

$$\begin{aligned} \pi_{\mu_{x_2}}(\mu_{x_2} | \mathbf{x}, \mathbf{z}) = & \frac{\sqrt{\tau_{x_2}^2 \sum z_i + \sigma_{x_2}^2}}{\sqrt{2\pi\tau_{x_2}\sigma_{x_2}}} \\ & \times \exp \left[-\frac{\tau_{x_2}^2 \sum z_i + \sigma_{x_2}^2}{2\tau_{x_2}^2 \sigma_{x_2}^2} \left(\mu_{x_2} - \frac{\bar{x}_{(2)}\tau_{x_2}^2 \sum z_i + \lambda_{x_2}\sigma_{x_2}^2}{\tau_{x_2}^2 \sum z_i + \sigma_{x_2}^2} \right)^2 \right], \end{aligned} \quad (3.4.17)$$

où

$$\bar{x}_{(1)} = \frac{\sum_{i=1}^m (1 - z_i) \log x_i}{\sum_{i=1}^m (1 - z_i)},$$

$$\bar{x}_{(2)} = \frac{\sum_{i=1}^m z_i \log x_i}{\sum_{i=1}^m z_i}.$$

Nous générons donc μ_{x_1} à l'aide d'une loi

$$N \left(\frac{\bar{x}_{(1)}\tau_{x_1}^2 \sum (1 - z_i) + \lambda_{x_1}\sigma_{x_1}^2}{\tau_{x_1}^2 \sum (1 - z_i) + \sigma_{x_1}^2}, \frac{\tau_{x_1}^2 \sigma_{x_1}^2}{\tau_{x_1}^2 \sum (1 - z_i) + \sigma_{x_1}^2} \right)$$

et μ_{x_2} à l'aide d'une loi

$$N \left(\frac{\bar{x}_{(2)}\tau_{x_2}^2 \sum z_i + \lambda_{x_2}\sigma_{x_2}^2}{\tau_{x_2}^2 \sum z_i + \sigma_{x_2}^2}, \frac{\tau_{x_2}^2 \sigma_{x_2}^2}{\tau_{x_2}^2 \sum z_i + \sigma_{x_2}^2} \right).$$

Ensuite trouvons $\pi_{\mathbf{z}}(\mathbf{z} | \mu_{\mathbf{x}}, \mathbf{x})$. Un problème surgit toutefois dans le fait que nous devons connaître \mathbf{z} pour générer $\mu_{\mathbf{x}}$ et vice-versa. Pour y remédier, nous estimerons μ_{x_1} par $\hat{\mu}_{x_1}$ et μ_{x_2} par $\hat{\mu}_{x_2}$ afin de générer \mathbf{z} , où $\hat{\mu}_{x_1}$ et $\hat{\mu}_{x_2}$ sont des estimateurs trouvés par l'algorithme EM. Ensuite nous pourrons générer $\mu_{\mathbf{x}}$ sans problème.

Nous posons donc

$$\begin{aligned} I_{\mathbf{Z}}(\mathbf{z}) &= \pi_{\mathbf{Z}}(\mathbf{z} \mid \hat{\mu}_{\mathbf{x}}, \mathbf{x}) \\ &= K_{\mathbf{X},3}(\mathbf{x})^{-1} h_{\mathbf{Z},1}(\mathbf{z} \mid \hat{\mu}_{x_1}, \mathbf{x}) h_{\mathbf{Z},2}(\mathbf{z} \mid \hat{\mu}_{x_2}, \mathbf{x}), \end{aligned} \quad (3.4.18)$$

où

$$h_{\mathbf{Z},1}(\mathbf{z} \mid \hat{\mu}_{x_1}, \mathbf{x}) = \left(\frac{1 - \epsilon_x}{\sigma_{x_1}} \right)^{\sum (1 - z_i)} \exp \left[-\frac{1}{2\sigma_{x_1}^2} \sum_{i=1}^m (\log x_i - \hat{\mu}_{x_1})^2 (1 - z_i) \right], \quad (3.4.19)$$

$$h_{\mathbf{Z},2}(\mathbf{z} \mid \hat{\mu}_{x_2}, \mathbf{x}) = \left(\frac{\epsilon_x}{\sigma_{x_2}} \right)^{\sum z_i} \exp \left[-\frac{1}{2\sigma_{x_2}^2} \sum_{i=1}^m (\log x_i - \hat{\mu}_{x_2})^2 z_i \right], \quad (3.4.20)$$

$$\begin{aligned} K_{\mathbf{X},3}(\mathbf{x}) &= \\ &= \prod_{i=1}^m \left(\frac{1 - \epsilon_x}{\sigma_{x_1}} \exp \left[-\frac{1}{2\sigma_{x_1}^2} (\log x_i - \hat{\mu}_{x_1})^2 \right] + \frac{\epsilon_x}{\sigma_{x_2}} \exp \left[-\frac{1}{2\sigma_{x_2}^2} (\log x_i - \hat{\mu}_{x_2})^2 \right] \right), \end{aligned} \quad (3.4.21)$$

comme nous l'avons démontré à la section 2.2.2. Ainsi nous générons z_i selon une loi

$$\text{Bin} \left(1, \frac{\frac{\epsilon_x}{\sigma_{x_2}} \exp \left[-\frac{1}{2\sigma_{x_2}^2} (\log x_i - \hat{\mu}_{x_2})^2 \right]}{\frac{1 - \epsilon_x}{\sigma_{x_1}} \exp \left[-\frac{1}{2\sigma_{x_1}^2} (\log x_i - \hat{\mu}_{x_1})^2 \right] + \frac{\epsilon_x}{\sigma_{x_2}} \exp \left[-\frac{1}{2\sigma_{x_2}^2} (\log x_i - \hat{\mu}_{x_2})^2 \right]} \right),$$

pour $i = 1, \dots, m$.

Il nous reste donc à calculer la fonction de poids $w(\boldsymbol{\theta})$, c'est-à-dire:

$$w(\mu_{\mathbf{x}}, \mu_{\mathbf{y}}, \mathbf{z}, \mathbf{w}) = w_{\mathbf{X}}(\mu_{\mathbf{x}}, \mathbf{z}) w_{\mathbf{Y}}(\mu_{\mathbf{y}}, \mathbf{w}),$$

où

$$\begin{aligned} w_{\mathbf{X}}(\mu_{\mathbf{x}}, \mathbf{z}) &= \frac{f_{\mathbf{X}}(\mathbf{x} \mid \mu_{\mathbf{x}}, \mathbf{z}) \pi_{\mu_{\mathbf{x}}}(\mu_{\mathbf{x}}) \pi_{\mathbf{Z}}(\mathbf{z})}{I_{\mu_{\mathbf{x}}}(\mu_{\mathbf{x}}) I_{\mathbf{Z}}(\mathbf{z})}, \\ w_{\mathbf{Y}}(\mu_{\mathbf{y}}, \mathbf{w}) &= \frac{f_{\mathbf{Y}}(\mathbf{y} \mid \mu_{\mathbf{y}}, \mathbf{w}) \pi_{\mu_{\mathbf{y}}}(\mu_{\mathbf{y}}) \pi_{\mathbf{W}}(\mathbf{w})}{I_{\mu_{\mathbf{y}}}(\mu_{\mathbf{y}}) I_{\mathbf{W}}(\mathbf{w})}. \end{aligned}$$

Développons $w_{\mathbf{X}}(\mu_{\mathbf{x}}, \mathbf{z})$. Il est à noter que $w_{\mathbf{Y}}(\mu_{\mathbf{y}}, \mathbf{w})$ est identique en changeant \mathbf{x} par \mathbf{y} et \mathbf{z} par \mathbf{w} . Alors nous pouvons écrire:

$$w_{\mathbf{X}}(\mu_{\mathbf{x}}, \mathbf{z}) = K_{\mathbf{X},1}(\mathbf{x})K_{\mathbf{X},2}(\mathbf{x})K_{\mathbf{X},3}(\mathbf{x})w_{\mathbf{X},1}(\hat{\mu}_{x_1}, \mathbf{z})w_{\mathbf{X},2}(\hat{\mu}_{x_2}, \mathbf{z})I_X(\mu_{x_1}, \mu_{x_2}),$$

où

$$w_{\mathbf{X},j}(\hat{\mu}_{x_j}, \mathbf{z}) = \frac{g_{\mathbf{X},j}(\mathbf{x} \mid \mu_{x_j}, \mathbf{z})\pi_{\mu_{x_j}}(\mu_{x_j})g_{\mathbf{Z},j}(\mathbf{z})}{\pi_{\mu_{x_j}}(\mu_{x_j} \mid \mathbf{x}, \mathbf{z})h_{\mathbf{Z},j}(\mathbf{z} \mid \hat{\mu}_{x_j}, \mathbf{x})}, \quad j = 1, 2,$$

en utilisant les équations (3.4.1), (3.4.4), (3.4.8), (3.4.15), et (3.4.18). Nous pouvons voir à l'annexe 1 que

$$\begin{aligned} w_{\mathbf{X},1}(\hat{\mu}_{x_1}, \mathbf{z}) &= \frac{\sigma_{x_1}}{\sqrt{\tau_{x_1}^2 \sum (1 - z_i) + \sigma_{x_1}^2}} \exp \left[\frac{1}{2\sigma_{x_1}^2} (\hat{\mu}_{x_1} - \bar{x}_{(1)})^2 \sum_{i=1}^m (1 - z_i) \right] \\ &\times \exp \left[\frac{-[\sum (1 - z_i)](\bar{x}_{(1)} - \lambda_{x_1})^2}{2(\tau_{x_1}^2 \sum (1 - z_i) + \sigma_{x_1}^2)} \right], \\ w_{\mathbf{X},2}(\hat{\mu}_{x_2}, \mathbf{z}) &= \frac{\sigma_{x_2}}{\sqrt{\tau_{x_2}^2 \sum z_i + \sigma_{x_2}^2}} \exp \left[\frac{1}{2\sigma_{x_2}^2} (\hat{\mu}_{x_2} - \bar{x}_{(2)})^2 \sum_{i=1}^m z_i \right] \\ &\times \exp \left[\frac{-(\sum z_i)(\bar{x}_{(2)} - \lambda_{x_2})^2}{2(\tau_{x_2}^2 \sum z_i + \sigma_{x_2}^2)} \right]. \end{aligned} \quad (3.4.22)$$

Comme $K_{\mathbf{X},1}(\mathbf{x})$, $K_{\mathbf{X},2}(\mathbf{x})$ et $K_{\mathbf{X},3}(\mathbf{x})$ sont des constantes qui se retrouvent dans l'intégrale du numérateur et dans celle du dénominateur, elles s'annuleront et nous pouvons par conséquent les ignorer dans le calcul de la fonction de poids.

Nous avons tous les éléments pour estimer $p(\mathbf{x}, \mathbf{y}) = \mathbb{P}[H_0 : \mathbb{E}(X) < \mathbb{E}(Y) \mid \mathbf{x}, \mathbf{y}]$. Voici l'algorithme à suivre:

1) Choisir les paramètres des lois *a priori* de $\mu_{x_1}, \mu_{x_2}, \mu_{y_1}$ et μ_{y_2} , c'est-à-dire $(\lambda_{x_1}, \tau_{x_1}^2)$, $(\lambda_{x_2}, \tau_{x_2}^2)$, $(\lambda_{y_1}, \tau_{y_1}^2)$ et $(\lambda_{y_2}, \tau_{y_2}^2)$.

2) Déterminer $\hat{\mu}_{x_1}, \hat{\mu}_{x_2}, \hat{\mu}_{y_1}, \hat{\mu}_{y_2}, \hat{\sigma}_{x_1}^2, \hat{\sigma}_{x_2}^2, \hat{\sigma}_{y_1}^2, \hat{\sigma}_{y_2}^2, \hat{\epsilon}_x$ et $\hat{\epsilon}_y$ par l'algorithme EM, puis fixer $\sigma_{x_1}^2 = \hat{\sigma}_{x_1}^2, \sigma_{x_2}^2 = \hat{\sigma}_{x_2}^2, \sigma_{y_1}^2 = \hat{\sigma}_{y_1}^2, \sigma_{y_2}^2 = \hat{\sigma}_{y_2}^2, \epsilon_x = \hat{\epsilon}_x$ et $\epsilon_y = \hat{\epsilon}_y$.

3) Générer z_i , pour $i = 1, \dots, m$, selon une loi

$$Bin \left(1, \frac{\frac{\epsilon_x}{\sigma_{x_2}} \exp \left[-\frac{1}{2\sigma_{x_2}^2} (\log x_i - \hat{\mu}_{x_2})^2 \right]}{\frac{1-\epsilon_x}{\sigma_{x_1}} \exp \left[-\frac{1}{2\sigma_{x_1}^2} (\log x_i - \hat{\mu}_{x_1})^2 \right] + \frac{\epsilon_x}{\sigma_{x_2}} \exp \left[-\frac{1}{2\sigma_{x_2}^2} (\log x_i - \hat{\mu}_{x_2})^2 \right]} \right) \quad (3.4.23)$$

4) Générer w_i , pour $i = 1, \dots, n$, selon une loi

$$Bin \left(1, \frac{\frac{\epsilon_y}{\sigma_{y_2}} \exp \left[-\frac{1}{2\sigma_{y_2}^2} (\log y_i - \hat{\mu}_{y_2})^2 \right]}{\frac{1-\epsilon_y}{\sigma_{y_1}} \exp \left[-\frac{1}{2\sigma_{y_1}^2} (\log y_i - \hat{\mu}_{y_1})^2 \right] + \frac{\epsilon_y}{\sigma_{y_2}} \exp \left[-\frac{1}{2\sigma_{y_2}^2} (\log y_i - \hat{\mu}_{y_2})^2 \right]} \right) \quad (3.4.24)$$

5) Générer μ_{x_1} selon une loi

$$N \left(\frac{\bar{x}_{(1)} \tau_{x_1}^2 \sum (1 - z_i) + \lambda_{x_1} \sigma_{x_1}^2}{\tau_{x_1}^2 \sum (1 - z_i) + \sigma_{x_1}^2}, \frac{\tau_{x_1}^2 \sigma_{x_1}^2}{\tau_{x_1}^2 \sum (1 - z_i) + \sigma_{x_1}^2} \right) \quad (3.4.25)$$

et générer μ_{x_2} selon une loi

$$N \left(\frac{\bar{x}_{(2)} \tau_{x_2}^2 \sum z_i + \lambda_{x_2} \sigma_{x_2}^2}{\tau_{x_2}^2 \sum z_i + \sigma_{x_2}^2}, \frac{\tau_{x_2}^2 \sigma_{x_2}^2}{\tau_{x_2}^2 \sum z_i + \sigma_{x_2}^2} \right). \quad (3.4.26)$$

6) Générer μ_{y_1} selon une loi

$$N \left(\frac{\bar{y}_{(1)} \tau_{y_1}^2 \sum (1 - w_i) + \lambda_{y_1} \sigma_{y_1}^2}{\tau_{y_1}^2 \sum (1 - w_i) + \sigma_{y_1}^2}, \frac{\tau_{y_1}^2 \sigma_{y_1}^2}{\tau_{y_1}^2 \sum (1 - w_i) + \sigma_{y_1}^2} \right) \quad (3.4.27)$$

et générer μ_{y_2} selon une loi

$$N \left(\frac{\bar{y}_{(2)} \tau_{y_2}^2 \sum w_i + \lambda_{y_2} \sigma_{y_2}^2}{\tau_{y_2}^2 \sum w_i + \sigma_{y_2}^2}, \frac{\tau_{y_2}^2 \sigma_{y_2}^2}{\tau_{y_2}^2 \sum w_i + \sigma_{y_2}^2} \right). \quad (3.4.28)$$

7) Calculer les fonctions indicatrices suivantes:

$$\begin{aligned}
 I_{XY}(\mu_x, \mu_y) &= \begin{cases} 1 & ; \text{ si } (1 - \epsilon_x)e^{\mu_{x_1} + \sigma_{x_1}^2/2} + \epsilon_x e^{\mu_{x_2} + \sigma_{x_2}^2/2} \\ & < (1 - \epsilon_y)e^{\mu_{y_1} + \sigma_{y_1}^2/2} + \epsilon_y e^{\mu_{y_2} + \sigma_{y_2}^2/2}, \\ 0 & ; \text{ sinon,} \end{cases} \\
 I_X(\mu_{x_1}, \mu_{x_2}) &= \begin{cases} 1 & ; \text{ si } e^{\mu_{x_1} + \sigma_{x_1}^2/2} < e^{\mu_{x_2} + \sigma_{x_2}^2/2}, \\ 0 & ; \text{ sinon,} \end{cases} \\
 I_Y(\mu_{y_1}, \mu_{y_2}) &= \begin{cases} 1 & ; \text{ si } e^{\mu_{y_1} + \sigma_{y_1}^2/2} < e^{\mu_{y_2} + \sigma_{y_2}^2/2}, \\ 0 & ; \text{ sinon.} \end{cases}
 \end{aligned}$$

8) Calculer

$$\begin{aligned}
 w_{\mathbf{X},1}(\hat{\mu}_{x_1}, \mathbf{z}) &= \frac{\sigma_{x_1}}{\sqrt{\tau_{x_1}^2 \sum (1 - z_i) + \sigma_{x_1}^2}} \exp \left[\frac{1}{2\sigma_{x_1}^2} (\hat{\mu}_{x_1} - \bar{x}_{(1)})^2 \sum_{i=1}^m (1 - z_i) \right] \\
 &\times \exp \left[\frac{-[\sum (1 - z_i)](\bar{x}_{(1)} - \lambda_{x_1})^2}{2(\tau_{x_1}^2 \sum (1 - z_i) + \sigma_{x_1}^2)} \right], \quad (3.4.29)
 \end{aligned}$$

$$\begin{aligned}
 w_{\mathbf{X},2}(\hat{\mu}_{x_2}, \mathbf{z}) &= \frac{\sigma_{x_2}}{\sqrt{\tau_{x_2}^2 \sum z_i + \sigma_{x_2}^2}} \exp \left[\frac{1}{2\sigma_{x_2}^2} (\hat{\mu}_{x_2} - \bar{x}_{(2)})^2 \sum_{i=1}^m z_i \right] \\
 &\times \exp \left[\frac{-(\sum z_i)(\bar{x}_{(2)} - \lambda_{x_2})^2}{2(\tau_{x_2}^2 \sum z_i + \sigma_{x_2}^2)} \right]. \quad (3.4.30)
 \end{aligned}$$

9) Calculer

$$\begin{aligned}
 w_{\mathbf{Y},1}(\hat{\mu}_{y_1}, \mathbf{w}) &= \frac{\sigma_{y_1}}{\sqrt{\tau_{y_1}^2 \sum (1 - w_i) + \sigma_{y_1}^2}} \exp \left[\frac{1}{2\sigma_{y_1}^2} (\hat{\mu}_{y_1} - \bar{y}_{(1)})^2 \sum_{i=1}^n (1 - w_i) \right] \\
 &\times \exp \left[\frac{-[\sum (1 - w_i)](\bar{y}_{(1)} - \lambda_{y_1})^2}{2(\tau_{y_1}^2 \sum (1 - w_i) + \sigma_{y_1}^2)} \right], \quad (3.4.31)
 \end{aligned}$$

$$\begin{aligned}
 w_{\mathbf{Y},2}(\hat{\mu}_{y_2}, \mathbf{w}) &= \frac{\sigma_{y_2}}{\sqrt{\tau_{y_2}^2 \sum w_i + \sigma_{y_2}^2}} \exp \left[\frac{1}{2\sigma_{y_2}^2} (\hat{\mu}_{y_2} - \bar{y}_{(2)})^2 \sum_{i=1}^n w_i \right] \\
 &\times \exp \left[\frac{-(\sum w_i)(\bar{y}_{(2)} - \lambda_{y_2})^2}{2(\tau_{y_2}^2 \sum w_i + \sigma_{y_2}^2)} \right]. \quad (3.4.32)
 \end{aligned}$$

10) Calculer

$$\begin{aligned}
w_j &= w_{\mathbf{X},1}(\hat{\mu}_{x_1}, \mathbf{z}) w_{\mathbf{X},2}(\hat{\mu}_{x_2}, \mathbf{z}) w_{\mathbf{Y},1}(\hat{\mu}_{y_1}, \mathbf{w}) w_{\mathbf{Y},2}(\hat{\mu}_{y_2}, \mathbf{w}) I_X(\mu_{x_1}, \mu_{x_2}) I_Y(\mu_{y_1}, \mu_{y_2}), \\
q_j &= I_{XY}(\mu_{\mathbf{x}}, \mu_{\mathbf{y}}).
\end{aligned}$$

11) Les étapes 3 à 10 représentent une simulation notée j . Refaire ces étapes un grand nombre de fois, disons S , afin de générer une suite w_1, \dots, w_S et une suite q_1, \dots, q_S .

12) Estimer la probabilité *a posteriori* de l'hypothèse H_0 par la formule suivante:

$$\mathbb{P}[H_0 : \mathbb{E}(X) < \mathbb{E}(Y) \mid \mathbf{x}, \mathbf{y}] = \frac{\sum_{j=1}^S q_j w_j}{\sum_{j=1}^S w_j}.$$

3.5. TEST D'HYPOTHÈSES BILATÉRAL AVEC POIDS ET VARIANCES CONNUS

Comme dans le test unilatéral de la section 3.4, les seuls paramètres considérés comme étant des variables aléatoires sont $\mu_{x_1}, \mu_{x_2}, \mu_{y_1}, \mu_{y_2}, \mathbf{z}$ et \mathbf{w} , les paramètres de poids (ϵ_x, ϵ_y) et de dispersion $(\sigma_{x_1}^2, \sigma_{x_2}^2, \sigma_{y_1}^2, \sigma_{y_2}^2)$ sont considérés connus. Aussi le modèle est le même et le test bilatéral sera semblable en de nombreux points au test unilatéral.

Le test d'hypothèses bilatéral est $H_0 : \mathbb{E}(X) = \mathbb{E}(Y)$ contre $H_1 : \mathbb{E}(X) \neq \mathbb{E}(Y)$, où $\mathbb{E}(X)$ et $\mathbb{E}(Y)$ sont données par les équations (3.4.12) et (3.4.13) respectivement.

Notons que les lois *a priori* sur \mathbf{z} et \mathbf{w} demeurent les mêmes que dans le test unilatéral, soit $z_i \sim \text{Bin}(1, \epsilon_x)$, $i = 1, \dots, m$ et $w_i \sim \text{Bin}(1, \epsilon_y)$, $i = 1, \dots, n$. Toutefois, nous ne pouvons garder les mêmes lois *a priori* continues pour $\mu_{x_1}, \mu_{x_2}, \mu_{y_1}$ et μ_{y_2} , sans quoi la probabilité *a priori* que H_0 survienne sera nulle puisque c'est une hypothèse discrète. Il faut donc utiliser une mesure de Dirac pour modéliser la probabilité *a priori* que H_0 survienne.

Notons aussi que $H_0 : \mathbb{E}(X) = \mathbb{E}(Y)$ est équivalent à $H_0 : \mu_{y_2} = \mu_{y_2}^*$, où $\mu_{y_2}^*$ est tel que

$$(1 - \epsilon_x)e^{\mu_{x_1} + \sigma_{x_1}^2/2} + \epsilon_x e^{\mu_{x_2} + \sigma_{x_2}^2/2} = (1 - \epsilon_y)e^{\mu_{y_1} + \sigma_{y_1}^2/2} + \epsilon_y e^{\mu_{y_2}^* + \sigma_{y_2}^2/2},$$

c'est-à-dire

$$\mu_{y_2}^* = \log \left[\frac{(1 - \epsilon_x)e^{\mu_{x_1} + \sigma_{x_1}^2/2} + \epsilon_x e^{\mu_{x_2} + \sigma_{x_2}^2/2} - (1 - \epsilon_y)e^{\mu_{y_1} + \sigma_{y_1}^2/2}}{\epsilon_y} \right] - \sigma_{y_2}^2/2. \quad (3.5.1)$$

Posons p_0 la probabilité *a priori* de l'hypothèse H_0 , c'est-à-dire

$$\begin{aligned} p_0 &= \mathbb{P}[H_0 : \mathbb{E}(X) = \mathbb{E}(Y)] \\ &= \mathbb{P}[\mu_{y_2} = \mu_{y_2}^*]. \end{aligned}$$

Ainsi nous définissons les lois a priori de $\mu_{x_1}, \mu_{x_2}, \mu_{y_1}, \mu_{y_2}$ de cette façon:

$$\pi_{\boldsymbol{\mu}}(\boldsymbol{\mu}_{\mathbf{x}}, \boldsymbol{\mu}_{\mathbf{y}}) = (1 - p_0)\pi_{\boldsymbol{\mu},1}(\boldsymbol{\mu}_{\mathbf{x}}, \boldsymbol{\mu}_{\mathbf{y}}) + p_0\pi_{\boldsymbol{\mu},2}(\boldsymbol{\mu}_{\mathbf{x}}, \boldsymbol{\mu}_{\mathbf{y}}),$$

où

$$\begin{aligned} \pi_{\boldsymbol{\mu},1}(\boldsymbol{\mu}_{\mathbf{x}}, \boldsymbol{\mu}_{\mathbf{y}}) &= K_{\mathbf{X},2}(\mathbf{x})K_{\mathbf{Y},2}(\mathbf{y})\pi_{\mu_{x_1}}(\mu_{x_1})\pi_{\mu_{x_2}}(\mu_{x_2})\pi_{\mu_{y_1}}(\mu_{y_1})\pi_{\mu_{y_2}}(\mu_{y_2}) \\ &\quad \times I_X(\mu_{x_1}, \mu_{x_2})I_Y(\mu_{y_1}, \mu_{y_2}), \\ \pi_{\boldsymbol{\mu},2}(\boldsymbol{\mu}_{\mathbf{x}}, \boldsymbol{\mu}_{\mathbf{y}}) &= K_{\mathbf{X},\mathbf{Y}}(\mathbf{x}, \mathbf{y})\pi_{\mu_{x_1}}(\mu_{x_1})\pi_{\mu_{x_2}}(\mu_{x_2})\pi_{\mu_{y_1}}(\mu_{y_1}) \\ &\quad \times I_{\mu_{y_2}^*} I_X(\mu_{x_1}, \mu_{x_2})I_Y^*(\mu_{y_1}, \mu_{y_2}^*), \end{aligned}$$

où $K_{\mathbf{X},2}(\mathbf{x})$, $K_{\mathbf{Y},2}(\mathbf{y})$, $\pi_{\mu_{x_1}}(\mu_{x_1})$, $\pi_{\mu_{x_2}}(\mu_{x_2})$, $\pi_{\mu_{y_1}}(\mu_{y_1})$, $\pi_{\mu_{y_2}}(\mu_{y_2})$, $I_X(\mu_{x_1}, \mu_{x_2})$ et $I_Y(\mu_{y_1}, \mu_{y_2})$ sont définies de la même façon que dans le test unilatéral selon les

équations (3.4.9) à (3.4.11) et

$$\begin{aligned}
I_{\mu_{y_2}^*} &= \begin{cases} 1 & ; \text{ si } \mu_{y_2} = \mu_{y_2}^*, \\ 0 & ; \text{ sinon,} \end{cases} \\
I_Y^*(\mu_{y_1}, \mu_{y_2}^*) &= \begin{cases} 1 & ; \text{ si } e^{\mu_{y_1} + \sigma_{y_1}^2/2} < e^{\mu_{y_2}^* + \sigma_{y_2}^2/2}, \\ 0 & ; \text{ sinon,} \end{cases} \\
K_{X,Y}(\mathbf{x}, \mathbf{y})^{-1} &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \pi_{\mu_{x_1}}(\mu_{x_1}) \pi_{\mu_{x_2}}(\mu_{x_2}) \pi_{\mu_{y_1}}(\mu_{y_1}) \\
&\quad \times I_{\mu_{y_2}^*} I_X(\mu_{x_1}, \mu_{x_2}) I_Y^*(\mu_{y_1}, \mu_{y_2}^*) d\mu_{x_1} d\mu_{x_2} d\mu_{y_1} d\mu_{y_2} \\
&= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \pi_{\mu_{x_1}}(\mu_{x_1}) \pi_{\mu_{x_2}}(\mu_{x_2}) \pi_{\mu_{y_1}}(\mu_{y_1}) \\
&\quad \times I_X(\mu_{x_1}, \mu_{x_2}) I_Y^*(\mu_{y_1}, \mu_{y_2}^*) d\mu_{x_1} d\mu_{x_2} d\mu_{y_1}.
\end{aligned}$$

Afin d'alléger la notation, nous notons $\boldsymbol{\mu} = (\mu_{\mathbf{x}}, \mu_{\mathbf{y}})$. Nous pouvons calculer $K_{X,2}(\mathbf{x})$, $K_{Y,2}(\mathbf{y})$ et $K_{X,Y}(\mathbf{x}, \mathbf{y})$ par la méthode de Monte Carlo. Contrairement au test unilatéral, il faudra les calculer puisque ces constantes ne s'annuleront pas en faisant le rapport d'intégrales.

Comme nous avons observé $\mathbf{x} = (x_1, \dots, x_m)^t$ et $\mathbf{y} = (y_1, \dots, y_n)^t$, nous sommes intéressés à calculer la probabilité *a posteriori* que l'hypothèse H_0 soit vraie, c'est-à-dire:

$$\begin{aligned}
p(\mathbf{x}, \mathbf{y}) &= \mathbb{P}[H_0 : \mathbb{E}(X) = \mathbb{E}(Y) \mid \mathbf{x}, \mathbf{y}] \\
&= \mathbb{P}[\mu_{y_2} = \mu_{y_2}^* \mid \mathbf{x}, \mathbf{y}] \\
&= \mathbb{E}^{\pi(\mu_{\mathbf{x}}, \mu_{\mathbf{y}}, \mathbf{z}, \mathbf{w} \mid \mathbf{x}, \mathbf{y})} (I_{\mu_{y_2}^*}) \\
&= \sum_{\mathbf{z}} \sum_{\mathbf{w}} \int_{\mu_{\mathbf{x}}} \int_{\mu_{\mathbf{y}}} I_{\mu_{y_2}^*} \pi(\mu_{\mathbf{x}}, \mu_{\mathbf{y}}, \mathbf{z}, \mathbf{w} \mid \mathbf{x}, \mathbf{y}) d\mu_{\mathbf{y}} d\mu_{\mathbf{x}} \\
&= \frac{\sum_{\mathbf{z}} \sum_{\mathbf{w}} \int_{\mu_{\mathbf{x}}} \int_{\mu_{\mathbf{y}}} I_{\mu_{y_2}^*} G d\mu_{\mathbf{y}} d\mu_{\mathbf{x}}}{\sum_{\mathbf{z}} \sum_{\mathbf{w}} \int_{\mu_{\mathbf{x}}} \int_{\mu_{\mathbf{y}}} G d\mu_{\mathbf{y}} d\mu_{\mathbf{x}}},
\end{aligned}$$

où

$$G = f_{X,Y}(\mathbf{x}, \mathbf{y} \mid \mu_{\mathbf{x}}, \mu_{\mathbf{y}}, \mathbf{z}, \mathbf{w}) \pi_{\boldsymbol{\mu}}(\mu_{\mathbf{x}}, \mu_{\mathbf{y}}) \pi_{\mathbf{Z}, \mathbf{W}}(\mathbf{z}, \mathbf{w}).$$

En remplaçant $\pi_{\boldsymbol{\mu}}(\mu_{\mathbf{x}}, \mu_{\mathbf{y}})$ par $(1-p_0)\pi_{\boldsymbol{\mu},1}(\mu_{\mathbf{x}}, \mu_{\mathbf{y}}) + p_0\pi_{\boldsymbol{\mu},2}(\mu_{\mathbf{x}}, \mu_{\mathbf{y}})$, nous trouvons:

$$p(\mathbf{x}, \mathbf{y}) = \frac{p_0 J_2(\mathbf{x}, \mathbf{y})}{(1-p_0) J_1(\mathbf{x}, \mathbf{y}) + p_0 J_2(\mathbf{x}, \mathbf{y})}$$

où

$$\begin{aligned}
J_1(\mathbf{x}, \mathbf{y}) &= \sum_{\mathbf{z}} \sum_{\mathbf{w}} \int_{\mu_{\mathbf{x}}} \int_{\mu_{\mathbf{y}}} f_{X,Y}(\mathbf{x}, \mathbf{y} \mid \mu_{\mathbf{x}}, \mu_{\mathbf{y}}, \mathbf{z}, \mathbf{w}) \pi_{\mu,1}(\mu_{\mathbf{x}}, \mu_{\mathbf{y}}) \pi_{\mathbf{Z}}(\mathbf{z}) \pi_{\mathbf{W}}(\mathbf{w}) \\
&\quad \times d\mu_{\mathbf{y}} d\mu_{\mathbf{x}}, \\
J_2(\mathbf{x}, \mathbf{y}) &= \sum_{\mathbf{z}} \sum_{\mathbf{w}} \int_{\mu_{\mathbf{x}}} \int_{\mu_{y_1}} f_{X,Y}(\mathbf{x}, \mathbf{y} \mid \mu_{\mathbf{x}}, \mu_{y_1}, \mu_{y_2}^*, \mathbf{z}, \mathbf{w}) \pi_{\mu,2}(\mu_{\mathbf{x}}, \mu_{\mathbf{y}}) \pi_{\mathbf{Z}}(\mathbf{z}) \\
&\quad \times \pi_{\mathbf{W}}(\mathbf{w}) d\mu_{y_1} d\mu_{\mathbf{x}}.
\end{aligned}$$

Alors en posant

$$R = \frac{J_2(\mathbf{x}, \mathbf{y})}{J_1(\mathbf{x}, \mathbf{y})},$$

nous obtenons

$$p(\mathbf{x}, \mathbf{y}) = \left(1 + \frac{1 - p_0}{p_0 R}\right)^{-1}.$$

Nous solutionnerons ce rapport d'intégrales, $R = \frac{J_2(\mathbf{x}, \mathbf{y})}{J_1(\mathbf{x}, \mathbf{y})}$, de façon similaire au test unilatéral en utilisant la méthode de Monte Carlo avec fonctions d'importance. Nous générons les variables $\mu_{x_1}, \mu_{x_2}, \mu_{y_1}, \mu_{y_2}, \mathbf{z}$ et \mathbf{w} par les mêmes fonctions d'importance que le test unilatéral, soient par les lois normales et binômiales que nous avons définies par les équations (3.4.23) à (3.4.28). Les fonctions de poids de l'intégrale $J_1(\mathbf{x}, \mathbf{y})$ sont identiques à celles du test unilatéral puisqu'il s'agit de la même intégrale que le dénominateur de $p(\mathbf{x}, \mathbf{y})$ du test unilatéral. Par contre, les fonctions de poids de $J_2(\mathbf{x}, \mathbf{y})$ changeront un peu.

Ainsi pour $J_1(\mathbf{x}, \mathbf{y})$, la fonction de poids est donnée par:

$$\begin{aligned}
w_{J_1(\mathbf{x}, \mathbf{y})}(\mu_{\mathbf{x}}, \mu_{\mathbf{y}}, \mathbf{z}, \mathbf{w}) &= w_{X,1}(\hat{\mu}_{x_1}, \mathbf{z}) w_{X,2}(\hat{\mu}_{x_2}, \mathbf{z}) w_{Y,1}(\hat{\mu}_{y_1}, \mathbf{w}) w_{Y,2}(\hat{\mu}_{y_2}, \mathbf{w}) \\
&\quad \times K_{X,1}(\mathbf{x}) K_{X,2}(\mathbf{x}) K_{X,3}(\mathbf{x}) K_{Y,1}(\mathbf{y}) K_{Y,2}(\mathbf{y}) K_{Y,3}(\mathbf{y}) \\
&\quad \times I_X(\mu_{x_1}, \mu_{x_2}) I_Y(\mu_{y_1}, \mu_{y_2}),
\end{aligned}$$

où $w_{X,1}(\hat{\mu}_{x_1}, \mathbf{z})$, $w_{X,2}(\hat{\mu}_{x_2}, \mathbf{z})$, $w_{Y,1}(\hat{\mu}_{y_1}, \mathbf{w})$ et $w_{Y,2}(\hat{\mu}_{y_2}, \mathbf{w})$ sont donnés par les équations (3.4.29) à (3.4.32), tandis que pour $J_2(\mathbf{x}, \mathbf{y})$, nous pouvons montrer

(détails à l'annexe 1) que la fonction de poids est donnée par:

$$\begin{aligned}
w_{J_2(\mathbf{x}, \mathbf{y})}(\mu_{\mathbf{x}}, \mu_{y_1}, \mathbf{z}, \mathbf{w}) &= w_{\mathbf{X},1}(\hat{\mu}_{x_1}, \mathbf{z}) w_{\mathbf{X},2}(\hat{\mu}_{x_2}, \mathbf{z}) w_{\mathbf{Y},1}(\hat{\mu}_{y_1}, \mathbf{w}) w'_{\mathbf{Y},2}(\hat{\mu}_{y_2}, \mu_{y_2}^*, \mathbf{w}) \\
&\times K_{\mathbf{X},1}(\mathbf{x}) K_{\mathbf{X},3}(\mathbf{x}) K_{\mathbf{Y},1}(\mathbf{y}) K_{\mathbf{Y},3}(\mathbf{y}) K_{\mathbf{X},\mathbf{Y}}(\mathbf{x}, \mathbf{y}) \\
&\times \mathbf{I}_X(\mu_{x_1}, \mu_{x_2}) \mathbf{I}_Y^*(\mu_{y_1}, \mu_{y_2}^*)
\end{aligned} \tag{3.5.2}$$

où

$$w'_{\mathbf{Y},2}(\hat{\mu}_{y_2}, \mu_{y_2}^*, \mathbf{w}) = \exp\left(-\frac{1}{2\sigma_{y_2}^2} \sum w_i [(\log y_i - \mu_{y_2}^*)^2 - (\log y_i - \hat{\mu}_{y_2})^2]\right).$$

Par la méthode de Monte Carlo avec fonctions d'importance, nous estimons R par:

$$R = \frac{J_2(\mathbf{x}, \mathbf{y})}{J_1(\mathbf{x}, \mathbf{y})} \approx \frac{\sum_{j=1}^S w_{J_2(\mathbf{x}, \mathbf{y})(j)}(\mu_{\mathbf{x}}, \mu_{y_1}, \mathbf{z}, \mathbf{w})}{\sum_{j=1}^S w_{J_1(\mathbf{x}, \mathbf{y})(j)}(\mu_{\mathbf{x}}, \mu_{\mathbf{y}}, \mathbf{z}, \mathbf{w})}.$$

Les constantes $K_{\mathbf{X},1}(\mathbf{x})$, $K_{\mathbf{X},3}(\mathbf{x})$, $K_{\mathbf{Y},1}(\mathbf{y})$ et $K_{\mathbf{Y},3}(\mathbf{y})$ se retrouvent au numérateur et au dénominateur du rapport R et nous pouvons par conséquent les ignorer dans le calcul de la fonction de poids.

Nous avons tous les éléments pour estimer $p(\mathbf{x}, \mathbf{y}) = \mathbb{P}[H_0 : \mathbb{E}(X) = \mathbb{E}(Y) \mid \mathbf{x}, \mathbf{y}]$. Voici l'algorithme à suivre:

- 1) Choisir les paramètres des lois *a priori* de $\mu_{x_1}, \mu_{x_2}, \mu_{y_1}$ et μ_{y_2} , c'est-à-dire $(\lambda_{x_1}, \tau_{x_1}^2)$, $(\lambda_{x_2}, \tau_{x_2}^2)$, $(\lambda_{y_1}, \tau_{y_1}^2)$ et $(\lambda_{y_2}, \tau_{y_2}^2)$.
- 2) Déterminer $\hat{\mu}_{x_1}, \hat{\mu}_{x_2}, \hat{\mu}_{y_1}, \hat{\mu}_{y_2}, \hat{\sigma}_{x_1}^2, \hat{\sigma}_{x_2}^2, \hat{\sigma}_{y_1}^2, \hat{\sigma}_{y_2}^2, \hat{\epsilon}_x$ et $\hat{\epsilon}_y$ par l'algorithme EM, puis fixer $\sigma_{x_1}^2 = \hat{\sigma}_{x_1}^2, \sigma_{x_2}^2 = \hat{\sigma}_{x_2}^2, \sigma_{y_1}^2 = \hat{\sigma}_{y_1}^2, \sigma_{y_2}^2 = \hat{\sigma}_{y_2}^2, \epsilon_x = \hat{\epsilon}_x$ et $\epsilon_y = \hat{\epsilon}_y$.
- 3) Déterminer $p_0 = \mathbb{P}[H_0 : \mathbb{E}(X) = \mathbb{E}(Y)]$ la probabilité *a priori* de l'hypothèse H_0 .
- 4) Calculer les constantes $K_{\mathbf{X},2}(\mathbf{x})$, $K_{\mathbf{Y},2}(\mathbf{y})$ et $K_{\mathbf{X},\mathbf{Y}}(\mathbf{x}, \mathbf{y})$ par la méthode de

Monte Carlo, où

$$\begin{aligned}
 K_{\mathbf{X},2}(\mathbf{x})^{-1} &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \pi_{\mu_{x_1}}(\mu_{x_1}) \pi_{\mu_{x_2}}(\mu_{x_2}) I_X(\mu_{x_1}, \mu_{x_2}) d\mu_{x_1} d\mu_{x_2}, \\
 K_{\mathbf{Y},2}(\mathbf{y})^{-1} &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \pi_{\mu_{y_1}}(\mu_{y_1}) \pi_{\mu_{y_2}}(\mu_{y_2}) I_Y(\mu_{y_1}, \mu_{y_2}) d\mu_{y_1} d\mu_{y_2}, \\
 K_{\mathbf{X},\mathbf{Y}}(\mathbf{x}, \mathbf{y})^{-1} &= \int_{\mu_{x_1}} \int_{\mu_{x_2}} \int_{\mu_{y_1}} \pi_{\mu_{x_1}}(\mu_{x_1}) \pi_{\mu_{x_2}}(\mu_{x_2}) \pi_{\mu_{y_1}}(\mu_{y_1}) I_X(\mu_{x_1}, \mu_{x_2}) \\
 &\quad \times I_Y^*(\mu_{y_1}, \mu_{y_2}^*) d\mu_{y_1} d\mu_{x_2} d\mu_{x_1}.
 \end{aligned}$$

5) Refaire les étapes 3 à 6 du test unilatéral, soit générer z_i , pour $i = 1, \dots, m$, w_i pour $i = 1, \dots, n$, μ_{x_1} , μ_{x_2} , μ_{y_1} et μ_{y_2} selon les équations (3.4.23) à (3.4.28).

6) Calculer

$$\mu_{y_2}^* = \log \left[\frac{(1 - \epsilon_x) e^{\mu_{x_1} + \sigma_{x_1}^2/2} + \epsilon_x e^{\mu_{x_2} + \sigma_{x_2}^2/2} - (1 - \epsilon_y) e^{\mu_{y_1} + \sigma_{y_1}^2/2}}{\epsilon_y} \right] - \sigma_{y_2}^2/2.$$

7) Calculer les fonctions indicatrices suivantes:

$$\begin{aligned}
 I_X(\mu_{x_1}, \mu_{x_2}) &= \begin{cases} 1 & ; \text{ si } e^{\mu_{x_1} + \sigma_{x_1}^2/2} < e^{\mu_{x_2} + \sigma_{x_2}^2/2}, \\ 0 & ; \text{ sinon,} \end{cases} \\
 I_Y(\mu_{y_1}, \mu_{y_2}) &= \begin{cases} 1 & ; \text{ si } e^{\mu_{y_1} + \sigma_{y_1}^2/2} < e^{\mu_{y_2} + \sigma_{y_2}^2/2}, \\ 0 & ; \text{ sinon,} \end{cases} \\
 I_Y^*(\mu_{y_1}, \mu_{y_2}^*) &= \begin{cases} 1 & ; \text{ si } e^{\mu_{y_1} + \sigma_{y_1}^2/2} < e^{\mu_{y_2}^* + \sigma_{y_2}^2/2}, \\ 0 & ; \text{ sinon.} \end{cases}
 \end{aligned}$$

8) Refaire les étapes 8 et 9 du test unilatéral, soit calculer les fonctions de poids $w_{\mathbf{X},1}(\hat{\mu}_{x_1}, \mathbf{z})$, $w_{\mathbf{X},2}(\hat{\mu}_{x_2}, \mathbf{z})$, $w_{\mathbf{Y},1}(\hat{\mu}_{y_1}, \mathbf{w})$ et $w_{\mathbf{Y},2}(\hat{\mu}_{y_2}, \mathbf{w})$ selon les équations (3.4.29) à (3.4.32).

9) Calculer

$$w'_{\mathbf{Y},2}(\hat{\mu}_{y_2}, \mu_{y_2}^*, \mathbf{w}) = \exp \left[-\frac{1}{2\sigma_{y_2}^2} \sum w_i [(\log y_i - \mu_{y_2}^*)^2 - (\log y_i - \hat{\mu}_{y_2})^2] \right].$$

10) Calculer

$$\begin{aligned} w'_j &= K_{\mathbf{X},2}(\mathbf{x})K_{\mathbf{Y},2}(\mathbf{y})w_{\mathbf{X},1}(\hat{\mu}_{x_1}, \mathbf{z})w_{\mathbf{X},2}(\hat{\mu}_{x_2}, \mathbf{z})w_{\mathbf{Y},1}(\hat{\mu}_{y_1}, \mathbf{w})w_{\mathbf{Y},2}(\hat{\mu}_{y_2}, \mathbf{w}) \\ &\quad \times I_X(\mu_{x_1}, \mu_{x_2})I_Y(\mu_{y_1}, \mu_{y_2}), \\ w''_j &= K_{\mathbf{X},\mathbf{Y}}(\mathbf{x}, \mathbf{y})w_{\mathbf{X},1}(\hat{\mu}_{x_1}, \mathbf{z})w_{\mathbf{X},2}(\hat{\mu}_{x_2}, \mathbf{z})w_{\mathbf{Y},1}(\hat{\mu}_{y_1}, \mathbf{w})w'_{\mathbf{Y},2}(\hat{\mu}_{y_2}, \mu_{y_2}^*, \mathbf{w}) \\ &\quad \times I_X(\mu_{x_1}, \mu_{x_2})I_Y^*(\mu_{y_1}, \mu_{y_2}^*). \end{aligned}$$

11) Les étapes 5 à 10 représentent une simulation notée j . Refaire ces étapes un grand nombre de fois, disons S , afin de générer une suite w'_1, \dots, w'_S et une suite w''_1, \dots, w''_S .

12) Estimer la probabilité par les formules suivantes:

$$R = \frac{\sum_{j=1}^S w''_j}{\sum_{j=1}^S w'_j}$$

et

$$p(\mathbf{x}, \mathbf{y}) = \mathbb{P}[H_0 : \mathbb{E}(X) = \mathbb{E}(Y) \mid \mathbf{x}, \mathbf{y}] = \left(1 + \frac{1 - p_0}{p_0 R}\right)^{-1}.$$

3.6. TEST D'HYPOTHÈSES UNILATÉRAL AVEC POIDS ET VARIANCES ALÉATOIRES

Les paramètres $\sigma_{x_1}^2, \sigma_{x_2}^2, \sigma_{y_1}^2, \sigma_{y_2}^2, \epsilon_x$ et ϵ_y sont maintenant considérés comme étant des variables aléatoires en plus des paramètres $\mu_{x_1}, \mu_{x_2}, \mu_{y_1}, \mu_{y_2}, \mathbf{z}$ et \mathbf{w} qui étaient déjà considérés aléatoires dans les sections 3.4 et 3.5.

Rappelons que les formules pour le traitement 2 seront identiques à celles du traitement 1 en changeant \mathbf{x} par \mathbf{y} et \mathbf{z} par \mathbf{w} .

Le modèle pour le traitement 1 reste le même mais s'écrit maintenant ainsi:

$$f_X(x_i \mid \mu_{\mathbf{x}}, \sigma_{\mathbf{x}}^2, \epsilon_x) = (1 - \epsilon_x)f_{X,1}(x_i \mid \mu_{x_1}, \sigma_{x_1}^2) + \epsilon_x f_{X,2}(x_i \mid \mu_{x_2}, \sigma_{x_2}^2), \quad i = 1, \dots, m,$$

que nous pouvons récrire comme:

$$f_X(x_i \mid \mu_{\mathbf{x}}, \sigma_{\mathbf{x}}^2, z_i) = f_{X,1}(x_i \mid \mu_{x_1}, \sigma_{x_1}^2)^{1-z_i} f_{X,2}(x_i \mid \mu_{x_2}, \sigma_{x_2}^2)^{z_i}, \quad i = 1, \dots, m,$$

où $z_i \mid \epsilon_x \sim Bin(1, \epsilon_x)$. Afin d'alléger la notation, nous écrirons $\mu_{\mathbf{x}}$ pour (μ_{x_1}, μ_{x_2}) , $\boldsymbol{\mu}$ pour $(\mu_{\mathbf{x}}, \mu_{\mathbf{y}})$, $\sigma_{\mathbf{x}}^2$ pour $(\sigma_{x_1}^2, \sigma_{x_2}^2)$, $\boldsymbol{\sigma}^2$ pour $(\sigma_{\mathbf{x}}^2, \sigma_{\mathbf{y}}^2)$ et ϵ pour (ϵ_x, ϵ_y) . Aussi,

$$f_{\mathbf{X}}(\mathbf{x} \mid \mu_{\mathbf{x}}, \sigma_{\mathbf{x}}^2, \mathbf{z}) = K_{\mathbf{X},1}(\mathbf{x})g_{\mathbf{X},1}(\mathbf{x} \mid \mu_{x_1}, \sigma_{x_1}^2, \mathbf{z})g_{\mathbf{X},2}(\mathbf{x} \mid \mu_{x_2}, \sigma_{x_2}^2, \mathbf{z}), \quad (3.6.1)$$

$$\pi_{\mathbf{Z}}(\mathbf{z} \mid \epsilon_x) = g_{\mathbf{Z},1}(\mathbf{z} \mid \epsilon_x)g_{\mathbf{Z},2}(\mathbf{z} \mid \epsilon_x), \quad (3.6.2)$$

tel que définies par les équations (3.4.1) à (3.4.7). Nous avons déjà défini à la section 3.4 des lois *a priori* pour les paramètres $\mu_{x_1}, \mu_{x_2}, \mu_{y_1}$ et μ_{y_2} . Nous devons maintenant en choisir pour les paramètres $\sigma_{x_1}^2, \sigma_{x_2}^2, \sigma_{y_1}^2, \sigma_{y_2}^2, \epsilon_x$ et ϵ_y . Nous avons vu à la section 3.2 que la famille de lois inverse-gamma pour les paramètres $\sigma_{x_1}^2$ et $\sigma_{x_2}^2$ est conjuguée pour la famille de lois log-normales. Ainsi nous choisissons

$$\sigma_{x_j}^2 \sim IG(\alpha_{x_j}, \beta_{x_j}), \quad j = 1, 2.$$

Notons que la fonction de densité d'une loi inverse-gamma est définie comme suit:

$$\pi_{\sigma_{x_j}^2}(\sigma_{x_j}^2) = \frac{\beta_{x_j}^{\alpha_{x_j}} \exp(-\beta_{x_j}/\sigma_{x_j}^2)}{\Gamma(\alpha_{x_j})(\sigma_{x_j}^2)^{\alpha_{x_j}+1}}, \quad j = 1, 2. \quad (3.6.3)$$

Comme nous voulons encore une fois que la moyenne de la première distribution log-normale soit plus petite que la deuxième, alors nous définissons la fonction de densité conjointe de $\mu_{x_1}, \mu_{x_2}, \sigma_{x_1}^2, \sigma_{x_2}^2$ comme suit:

$$\pi(\mu_{x_1}, \mu_{x_2}, \sigma_{x_1}^2, \sigma_{x_2}^2) = K_{\mathbf{X},2}(\mathbf{x})\pi_{\mu_{x_1}}(\mu_{x_1})\pi_{\mu_{x_2}}(\mu_{x_2})\pi_{\sigma_{x_1}^2}(\sigma_{x_1}^2)\pi_{\sigma_{x_2}^2}(\sigma_{x_2}^2)I_X(\mu_{\mathbf{x}}, \sigma_{\mathbf{x}}^2), \quad (3.6.4)$$

où

$$K_{\mathbf{X},2}(\mathbf{x})^{-1} = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \int_0^{\infty} \int_0^{\infty} \pi_{\mu_{x_1}}(\mu_{x_1})\pi_{\mu_{x_2}}(\mu_{x_2})\pi_{\sigma_{x_1}^2}(\sigma_{x_1}^2)\pi_{\sigma_{x_2}^2}(\sigma_{x_2}^2)I_X(\mu_{\mathbf{x}}, \sigma_{\mathbf{x}}^2) \\ \times d\sigma_{x_1}^2 d\sigma_{x_2}^2 d\mu_{x_1} d\mu_{x_2}, \quad (3.6.5)$$

$$I_X(\mu_{\mathbf{x}}, \sigma_{\mathbf{x}}^2) = \begin{cases} 1 & ; \text{ si } \mathbb{E}_1(X) < \mathbb{E}_2(X), \\ 0 & ; \text{ sinon,} \end{cases} \quad (3.6.6)$$

et $\pi_{\mu_{x_1}}(\mu_{x_1}), \pi_{\mu_{x_2}}(\mu_{x_2}), \pi_{\sigma_{x_1}^2}(\sigma_{x_1}^2)$ et $\pi_{\sigma_{x_2}^2}(\sigma_{x_2}^2)$ sont telles que déjà définies par les équations (3.4.10) et (3.6.3).

Aussi nous avons vu à la section 3.2 que la famille de lois bêta pour le paramètre ϵ_x est conjuguée pour la famille de lois binômiales ($z_i \mid \epsilon_x \sim Bin(1, \epsilon_x)$). Nous choisissons donc $\epsilon_x \sim Be(c_x, d_x)$. La fonction de densité est définie comme suit:

$$\pi_{\epsilon_x}(\epsilon_x) = K_{\mathbf{X},4}(\mathbf{x})g_{\epsilon_x,1}(\epsilon_x)g_{\epsilon_x,2}(\epsilon_x), \quad (3.6.7)$$

où

$$\begin{aligned} K_{\mathbf{X},4}(\mathbf{x}) &= \frac{\Gamma(c_x + d_x)}{\Gamma(c_x)\Gamma(d_x)}, \\ g_{\epsilon_x,1}(\epsilon_x) &= (1 - \epsilon_x)^{d_x - 1}, \\ g_{\epsilon_x,2}(\epsilon_x) &= (\epsilon_x)^{c_x - 1}. \end{aligned}$$

Rappelons que nous avons les mêmes résultats pour le traitement 2 en changeant \mathbf{x} par \mathbf{y} et \mathbf{z} par \mathbf{w} .

Le test d'hypothèses unilatéral est le suivant:

$$H_0 : \mathbb{E}(X) < \mathbb{E}(Y) \text{ contre } H_1 : \mathbb{E}(X) \geq \mathbb{E}(Y),$$

où $\mathbb{E}(X)$ et $\mathbb{E}(Y)$ sont données par les équations (3.4.12) et (3.4.13) respectivement.

Comme nous avons observé $\mathbf{x} = (x_1, \dots, x_m)^t$ et $\mathbf{y} = (y_1, \dots, y_n)^t$, nous sommes intéressés à calculer, tel que vu à la section 3.1, la probabilité *a posteriori* que l'hypothèse H_0 soit vraie, c'est-à-dire:

$$\begin{aligned} p(\mathbf{x}, \mathbf{y}) &= \mathbb{P}[H_0 : \mathbb{E}(X) < \mathbb{E}(Y) \mid \mathbf{x}, \mathbf{y}] \\ &= \mathbb{E}^{\pi(\mu_{\mathbf{x}}, \mu_{\mathbf{y}}, \sigma_{\mathbf{x}}^2, \sigma_{\mathbf{y}}^2, \epsilon_x, \epsilon_y, \mathbf{z}, \mathbf{w} \mid \mathbf{x}, \mathbf{y})} (I_{XY}(\boldsymbol{\mu}, \boldsymbol{\sigma}^2, \boldsymbol{\epsilon})) \\ &= \sum_{\mathbf{z}} \sum_{\mathbf{w}} \int_{\boldsymbol{\mu}} \int_{\boldsymbol{\sigma}^2} \int_{\boldsymbol{\epsilon}} I_{XY}(\boldsymbol{\mu}, \boldsymbol{\sigma}^2, \boldsymbol{\epsilon}) \pi(\boldsymbol{\mu}, \boldsymbol{\sigma}^2, \boldsymbol{\epsilon}, \mathbf{z}, \mathbf{w} \mid \mathbf{x}, \mathbf{y}) d\boldsymbol{\epsilon} d\boldsymbol{\sigma}^2 d\boldsymbol{\mu} \end{aligned}$$

où

$$I_{XY}(\boldsymbol{\mu}, \boldsymbol{\sigma}^2, \boldsymbol{\epsilon}) = \begin{cases} 1 & ; \text{ si } \mathbb{E}(X) < \mathbb{E}(Y), \\ 0 & ; \text{ sinon,} \end{cases}$$

et

$$\pi(\mu_{\mathbf{x}}, \mu_{\mathbf{y}}, \sigma_{\mathbf{x}}^2, \sigma_{\mathbf{y}}^2, \epsilon_x, \epsilon_y, \mathbf{z}, \mathbf{w} \mid \mathbf{x}, \mathbf{y}) = \frac{f_{\mathbf{X}, \mathbf{Y}}(\mathbf{x}, \mathbf{y} \mid \boldsymbol{\mu}, \boldsymbol{\sigma}^2, \boldsymbol{\epsilon}, \mathbf{z}, \mathbf{w}) \pi(\boldsymbol{\mu}, \boldsymbol{\sigma}^2, \boldsymbol{\epsilon}, \mathbf{z}, \mathbf{w})}{\sum_{\mathbf{z}} \sum_{\mathbf{w}} \int_{\boldsymbol{\mu}} \int_{\boldsymbol{\sigma}^2} \int_{\boldsymbol{\epsilon}} f_{\mathbf{X}, \mathbf{Y}}(\mathbf{x}, \mathbf{y} \mid \boldsymbol{\mu}, \boldsymbol{\sigma}^2, \boldsymbol{\epsilon}, \mathbf{z}, \mathbf{w}) \pi(\boldsymbol{\mu}, \boldsymbol{\sigma}^2, \boldsymbol{\epsilon}, \mathbf{z}, \mathbf{w}) d\boldsymbol{\epsilon} d\boldsymbol{\sigma}^2 d\boldsymbol{\mu}}.$$

Nous résoudrons cette intégrale par la méthode de simulation Monte Carlo avec fonctions d'importance telle que déjà décrite à la section 3.4. Ainsi

$$\begin{aligned} \boldsymbol{\theta} & \text{ correspond à } \mu_{\mathbf{x}}, \mu_{\mathbf{y}}, \sigma_{\mathbf{x}}^2, \sigma_{\mathbf{y}}^2, \epsilon_x, \epsilon_y, \mathbf{z}, \mathbf{w}, \\ q(\boldsymbol{\theta}) & \text{ correspond à } I_{\mathbf{X}\mathbf{Y}}(\boldsymbol{\mu}, \boldsymbol{\sigma}^2, \boldsymbol{\epsilon}), \\ I(\boldsymbol{\theta}) & \text{ correspond à } I_{\mu_{\mathbf{x}}}(\mu_{\mathbf{x}}) I_{\mu_{\mathbf{y}}}(\mu_{\mathbf{y}}) I_{\mathbf{Z}}(\mathbf{z}) I_{\mathbf{W}}(\mathbf{w}) I_{\sigma_{\mathbf{x}}^2}(\sigma_{\mathbf{x}}^2) I_{\sigma_{\mathbf{y}}^2}(\sigma_{\mathbf{y}}^2) I_{\epsilon_x}(\epsilon_x) I_{\epsilon_y}(\epsilon_y). \end{aligned}$$

Tel que mentionné à la section 3.4, nous essaierons de choisir les fonctions d'importance de façon à être le plus près possible des distributions *a posteriori*. Il faudra encore une fois fixer certaines variables afin d'éviter de tourner en rond. Ainsi, nous générerons les paramètres dans l'ordre: $\epsilon_x, \mathbf{z}, \sigma_{\mathbf{x}}^2$ puis finalement $\mu_{\mathbf{x}}$ en remplaçant les paramètres non générés par leur estimateur lorsque nécessaire. Ces estimateurs sont calculés par l'algorithme EM et sont identifiés par un chapeau. Les résultats sont les mêmes pour le traitement 2 en changeant \mathbf{x} par \mathbf{y} et \mathbf{z} par \mathbf{w} . Trouvons d'abord $I_{\epsilon_x}(\epsilon_x)$:

$$\begin{aligned} \pi_{\epsilon_x}(\epsilon_x \mid \mu_{\mathbf{x}}, \sigma_{\mathbf{x}}^2, \mathbf{x}, \mathbf{z}) & \propto f_{\mathbf{X}}(\mathbf{x} \mid \mathbf{z}, \mu_{\mathbf{x}}, \sigma_{\mathbf{x}}^2) \pi_{\mathbf{Z}}(\mathbf{z} \mid \epsilon_x) \pi_{\epsilon_x}(\epsilon_x) \\ & \propto \pi_{\mathbf{Z}}(\mathbf{z} \mid \epsilon_x) \pi_{\epsilon_x}(\epsilon_x) \\ & \propto (1 - \epsilon_x)^{m - \sum z_i} (\epsilon_x)^{\sum z_i} (1 - \epsilon_x)^{d_x - 1} (\epsilon_x)^{c_x - 1}. \end{aligned}$$

Alors

$$\pi_{\epsilon_x}(\epsilon_x \mid \mathbf{z}) = \frac{\Gamma(m + c_x + d_x)}{\Gamma(c_x + \sum z_i) \Gamma(d_x + m - \sum z_i)} (1 - \epsilon_x)^{m - \sum z_i + d_x - 1} (\epsilon_x)^{\sum z_i + c_x - 1}.$$

Or comme \mathbf{z} n'est pas encore généré, nous estimons $\sum z_i$ par $m\hat{\epsilon}_x$, où $\hat{\epsilon}_x$ est l'estimateur EM de ϵ_x . Alors posons

$$\begin{aligned} I_{\epsilon_x}(\epsilon_x) & = \pi_{\epsilon_x}(\epsilon_x \mid \hat{\epsilon}_x) \\ & = K_{\mathbf{X}, 5}(\mathbf{x}) h_{\epsilon_x, 1}(\epsilon_x \mid \hat{\epsilon}_x) h_{\epsilon_x, 2}(\epsilon_x \mid \hat{\epsilon}_x), \end{aligned} \quad (3.6.8)$$

où

$$\begin{aligned} K_{\mathbf{X},5}(\mathbf{x}) &= \frac{\Gamma(m + c_x + d_x)}{\Gamma(c_x + m\hat{\epsilon}_x)\Gamma(d_x + m(1 - \hat{\epsilon}_x))}, \\ h_{\epsilon_x,1}(\epsilon_x | \hat{\epsilon}_x) &= (1 - \epsilon_x)^{m(1-\hat{\epsilon}_x)+d_x-1}, \\ h_{\epsilon_x,2}(\epsilon_x | \hat{\epsilon}_x) &= (\epsilon_x)^{m\hat{\epsilon}_x+c_x-1}. \end{aligned}$$

Nous générons donc ϵ_x selon une loi

$$Be(c_x + m\hat{\epsilon}_x, d_x + m(1 - \hat{\epsilon}_x)).$$

Nous posons ensuite $I_{\mathbf{Z}}(\mathbf{z})$ comme étant la distribution *a posteriori* $\pi_{\mathbf{Z}}(\mathbf{z} | \mu_{\mathbf{x}}, \sigma_{\mathbf{x}}^2, \epsilon_x, \mathbf{x})$ trouvée à la section 3.4. Toutefois, comme $\mu_{\mathbf{x}}$ et $\sigma_{\mathbf{x}}^2$ ne sont pas encore générées, nous estimons $\mu_{\mathbf{x}}$ par $\hat{\mu}_{\mathbf{x}}$ et $\sigma_{\mathbf{x}}^2$ par $\hat{\sigma}_{\mathbf{x}}^2$, soient leurs estimateurs EM. Alors posons

$$\begin{aligned} I_{\mathbf{Z}}(\mathbf{z}) &= \pi_{\mathbf{Z}}(\mathbf{z} | \hat{\mu}_{\mathbf{x}}, \hat{\sigma}_{\mathbf{x}}^2, \epsilon_x, \mathbf{x}) \\ &= K_{\mathbf{X},3}(\epsilon_x, \mathbf{x})^{-1} h_{\mathbf{Z},1}(\mathbf{z} | \hat{\mu}_{x_1}, \hat{\sigma}_{x_1}^2, \epsilon_x, \mathbf{x}) h_{\mathbf{Z},2}(\mathbf{z} | \hat{\mu}_{x_2}, \hat{\sigma}_{x_2}^2, \epsilon_x, \mathbf{x}) \end{aligned} \tag{3.6.9}$$

tel que défini par les équations (3.4.18) à (3.4.21) en remplaçant μ_x par $\hat{\mu}_x$ et σ_x^2 par $\hat{\sigma}_x^2$.

Nous générons donc z_i selon une loi

$$Bin \left(1, \frac{\frac{\epsilon_x}{\hat{\sigma}_{x_2}} \exp \left[-\frac{1}{2\hat{\sigma}_{x_2}^2} (\log x_i - \hat{\mu}_{x_2})^2 \right]}{\frac{1-\epsilon_x}{\hat{\sigma}_{x_1}} \exp \left[-\frac{1}{2\hat{\sigma}_{x_1}^2} (\log x_i - \hat{\mu}_{x_1})^2 \right] + \frac{\epsilon_x}{\hat{\sigma}_{x_2}} \exp \left[-\frac{1}{2\hat{\sigma}_{x_2}^2} (\log x_i - \hat{\mu}_{x_2})^2 \right]} \right)$$

pour $i = 1, \dots, m$.

Trouvons maintenant $I_{\sigma_{\mathbf{x}}^2}(\sigma_{\mathbf{x}}^2)$:

$$\begin{aligned}
\pi_{\sigma_{\mathbf{x}}^2}(\sigma_{\mathbf{x}}^2 \mid \mu_{\mathbf{x}}, \epsilon_x, \mathbf{x}, \mathbf{z}) &\propto f_{\mathbf{X}}(\mathbf{x} \mid \mathbf{z}, \mu_{\mathbf{x}}, \sigma_{\mathbf{x}}^2) \pi_{\mathbf{Z}}(\mathbf{z} \mid \epsilon_x) \pi_{\sigma_{\mathbf{x}}^2}(\sigma_{\mathbf{x}}^2 \mid \mu_{\mathbf{x}}) \\
&\propto f_{\mathbf{X}}(\mathbf{x} \mid \mathbf{z}, \mu_{\mathbf{x}}, \sigma_{\mathbf{x}}^2) \pi_{\sigma_{x_1}^2}(\sigma_{x_1}^2) \pi_{\sigma_{x_2}^2}(\sigma_{x_2}^2) I_X(\mu_{\mathbf{x}}, \sigma_{\mathbf{x}}^2) \\
&\propto (\sigma_{x_1}^2)^{-(m-\sum z_i)/2} \exp \left[-\frac{1}{2\sigma_{x_1}^2} \sum (\log x_i - \mu_{x_1})^2 (1 - z_i) \right] \\
&\times \frac{1}{(\sigma_{x_1}^2)^{\alpha_{x_1}+1}} \exp \left[-\frac{\beta_{x_1}}{\sigma_{x_1}^2} \right] \\
&\times (\sigma_{x_2}^2)^{-\sum z_i/2} \exp \left[-\frac{1}{2\sigma_{x_2}^2} \sum (\log x_i - \mu_{x_2})^2 z_i \right] \\
&\times \frac{1}{(\sigma_{x_2}^2)^{\alpha_{x_2}+1}} \exp \left[-\frac{\beta_{x_2}}{\sigma_{x_2}^2} \right] I_X(\mu_{\mathbf{x}}, \sigma_{\mathbf{x}}^2).
\end{aligned}$$

Toutefois, comme $\mu_{\mathbf{x}}$ n'est pas encore générée, nous l'estimons par $\hat{\mu}_{\mathbf{x}}$ son estimateur EM. Afin de simplifier les calculs, nous laissons tomber le terme $I_X(\mu_{\mathbf{x}}, \sigma_{\mathbf{x}}^2)$ afin de générer sans cette contrainte. Alors posons

$$\begin{aligned}
I_{\sigma_{x_2}^2}(\sigma_{x_2}^2) &= \pi_{\sigma_{x_2}^2}(\sigma_{x_2}^2 \mid \hat{\mu}_{x_2}, \mathbf{x}, \mathbf{z}) \tag{3.6.10} \\
&= \frac{\left[\frac{1}{2} \sum (\log x_i - \hat{\mu}_{x_2})^2 z_i + \beta_{x_2} \right]^{\sum z_i + \alpha_{x_2}}}{\Gamma(\frac{\sum z_i}{2} + \alpha_{x_2}) (\sigma_{x_2}^2)^{\sum z_i/2 + 1 + \alpha_{x_2}}} \\
&\times \exp \left[-\frac{1}{\sigma_{x_2}^2} \left(\frac{1}{2} \sum (\log x_i - \hat{\mu}_{x_2})^2 z_i + \beta_{x_2} \right) \right],
\end{aligned}$$

et $I_{\sigma_{x_1}^2}(\sigma_{x_1}^2) = \pi_{\sigma_{x_1}^2}(\sigma_{x_1}^2 \mid \hat{\mu}_{x_1}, \mathbf{x}, \mathbf{z})$ est identique en remplaçant x_2 par x_1 et z_i par $1 - z_i$.

Nous générons donc $\sigma_{x_1}^2$ selon une loi

$$IG \left(\frac{\sum (1 - z_i)}{2} + \alpha_{x_1}, \frac{1}{2} \sum (\log x_i - \hat{\mu}_{x_1})^2 (1 - z_i) + \beta_{x_1} \right)$$

et $\sigma_{x_2}^2$ selon une loi

$$IG \left(\frac{\sum z_i}{2} + \alpha_{x_2}, \frac{1}{2} \sum (\log x_i - \hat{\mu}_{x_2})^2 z_i + \beta_{x_2} \right).$$

Nous voulons finalement trouver $I_{\mu_{\mathbf{x}}}(\mu_{\mathbf{x}})$ que nous choisissons de la même façon qu'à la section 3.4, soit:

$$I_{\mu_{\mathbf{x}}}(\mu_{\mathbf{x}}) = \pi_{\mu_{x_1}}(\mu_{x_1} \mid \sigma_{x_1}^2, \mathbf{x}, \mathbf{z}) \pi_{\mu_{x_2}}(\mu_{x_2} \mid \sigma_{x_2}^2, \mathbf{x}, \mathbf{z}), \tag{3.6.11}$$

tel que définie par les équations (3.4.15) à (3.4.17). Nous générons donc μ_{x_1} selon une loi

$$N\left(\frac{\bar{x}_{(1)}\tau_{x_1}^2 \sum(1-z_i) + \lambda_{x_1}\sigma_{x_1}^2}{\tau_{x_1}^2 \sum(1-z_i) + \sigma_{x_1}^2}, \frac{\tau_{x_1}^2 \sigma_{x_1}^2}{\tau_{x_1}^2 \sum(1-z_i) + \sigma_{x_1}^2}\right)$$

et μ_{x_2} selon une loi

$$N\left(\frac{\bar{x}_{(2)}\tau_{x_2}^2 \sum z_i + \lambda_{x_2}\sigma_{x_2}^2}{\tau_{x_2}^2 \sum z_i + \sigma_{x_2}^2}, \frac{\tau_{x_2}^2 \sigma_{x_2}^2}{\tau_{x_2}^2 \sum z_i + \sigma_{x_2}^2}\right).$$

Alors, selon la méthode de simulations Monte Carlo avec fonctions d'importance, il faut calculer la fonction de poids $w_{\mathbf{X}}(\mu_{\mathbf{x}}, \sigma_{\mathbf{x}}^2, \epsilon_x, \mathbf{z})w_{\mathbf{Y}}(\mu_{\mathbf{y}}, \sigma_{\mathbf{y}}^2, \epsilon_y, \mathbf{w})$, où

$$w_{\mathbf{X}}(\mu_{\mathbf{x}}, \sigma_{\mathbf{x}}^2, \epsilon_x, \mathbf{z}) = \frac{f_{\mathbf{X}}(\mathbf{x} \mid \mu_{\mathbf{x}}, \sigma_{\mathbf{x}}^2, \mathbf{z})\pi(\mu_{\mathbf{x}}, \sigma_{\mathbf{x}}^2)\pi_{\epsilon_x}(\epsilon_x)\pi_{\mathbf{Z}}(\mathbf{z} \mid \epsilon_x)}{I_{\mu_{\mathbf{x}}}(\mu_{\mathbf{x}})I_{\sigma_{\mathbf{x}}^2}(\sigma_{\mathbf{x}}^2)I_{\epsilon_x}(\epsilon_x)I_{\mathbf{Z}}(\mathbf{z})},$$

$$w_{\mathbf{Y}}(\mu_{\mathbf{y}}, \sigma_{\mathbf{y}}^2, \epsilon_y, \mathbf{w}) = \frac{f_{\mathbf{Y}}(\mathbf{y} \mid \mu_{\mathbf{y}}, \sigma_{\mathbf{y}}^2, \mathbf{w})\pi(\mu_{\mathbf{y}}, \sigma_{\mathbf{y}}^2)\pi_{\epsilon_y}(\epsilon_y)\pi_{\mathbf{W}}(\mathbf{w} \mid \epsilon_y)}{I_{\mu_{\mathbf{y}}}(\mu_{\mathbf{y}})I_{\sigma_{\mathbf{y}}^2}(\sigma_{\mathbf{y}}^2)I_{\epsilon_y}(\epsilon_y)I_{\mathbf{W}}(\mathbf{w})}.$$

Développons $w_{\mathbf{X}}(\mu_{\mathbf{x}}, \sigma_{\mathbf{x}}^2, \epsilon_x, \mathbf{z})$. La fonction de poids pour $w_{\mathbf{Y}}(\mu_{\mathbf{y}}, \sigma_{\mathbf{y}}^2, \epsilon_y, \mathbf{w})$ est identique en changeant \mathbf{x} par \mathbf{y} et \mathbf{z} par \mathbf{w} . Alors nous pouvons écrire, en utilisant les équations (3.6.1) à (3.6.11),

$$w_{\mathbf{X}}(\mu_{\mathbf{x}}, \sigma_{\mathbf{x}}^2, \epsilon_x, \mathbf{z}) = w_{\mathbf{X},1}(\mu_{x_1}, \sigma_{x_1}^2, \epsilon_x, \mathbf{z})w_{\mathbf{X},2}(\mu_{x_2}, \sigma_{x_2}^2, \epsilon_x, \mathbf{z}) \\ \times K_{\mathbf{X},1}(\mathbf{x})K_{\mathbf{X},2}(\mathbf{x})K_{\mathbf{X},3}(\epsilon_x, \mathbf{x})K_{\mathbf{X},4}(\mathbf{x})K_{\mathbf{X},5}(\mathbf{x})^{-1}I_{\mathbf{X}}(\mu_{\mathbf{x}}, \sigma_{\mathbf{x}}^2),$$

où pour $j = 1, 2$, nous avons

$$w_{\mathbf{X},j}(\mu_{x_j}, \sigma_{x_j}^2, \epsilon_x, \mathbf{z}) = \frac{g_{\mathbf{X},j}(\mathbf{x} \mid \mu_{x_j}, \sigma_{x_j}^2, \mathbf{z})\pi_{\mu_{x_j}}(\mu_{x_j})\pi_{\sigma_{x_j}^2}(\sigma_{x_j}^2)g_{\epsilon_x,j}(\epsilon_x)g_{\mathbf{Z},j}(\mathbf{z} \mid \epsilon_x)}{\pi_{\mu_{x_j}}(\mu_{x_j} \mid \sigma_{x_j}^2, \mathbf{x}, \mathbf{z})\pi_{\sigma_{x_j}^2}(\sigma_{x_j}^2 \mid \hat{\mu}_{x_j}, \mathbf{x}, \mathbf{z})h_{\epsilon_x,j}(\epsilon_x \mid \hat{\epsilon}_x)h_{\mathbf{Z},j}(\mathbf{z} \mid \hat{\mu}_{x_j}, \hat{\sigma}_{x_j}^2, \epsilon_x, \mathbf{x})},$$

Nous pouvons voir à l'annexe 1 que

$$\begin{aligned}
& w_{\mathbf{X},2}(\mu_{x_2}, \sigma_{x_2}^2, \epsilon_x, \mathbf{z}) \\
&= \exp \left[\frac{1}{2\sigma_{x_2}^2} (\hat{\mu}_{x_2} - \bar{x}_{(2)})^2 \sum_{i=1}^m z_i \right] \exp \left[-\frac{(\sum_i z_i)(\bar{x}_{(2)} - \lambda_{x_2})^2}{2(\tau_{x_2}^2 \sum_i z_i + \sigma_{x_2}^2)} \right] \\
&\times \frac{\sigma_{x_2}}{\sqrt{\tau_{x_2}^2 \sum_i z_i + \sigma_{x_2}^2}} \cdot \frac{\beta_{x_2}^{\alpha_{x_2}}}{\Gamma(\alpha_{x_2})} \cdot \frac{\Gamma(\frac{\sum_i z_i}{2} + \alpha_{x_2})}{\left[\frac{1}{2} \sum (\log x_i - \hat{\mu}_{x_2})^2 z_i + \beta_{x_2} \right]^{\frac{\sum_i z_i}{2} + \alpha_{x_2}}} \\
&\times \left(\frac{\epsilon_x^{m\hat{\epsilon}_x}}{(\hat{\sigma}_{x_2})^{\sum z_i}} \exp \left[-\frac{1}{2\hat{\sigma}_{x_2}^2} (\log x_i - \hat{\mu}_{x_2})^2 z_i \right] \right)^{-1}. \tag{3.6.12}
\end{aligned}$$

Comme $K_{\mathbf{X},1}(\mathbf{x})$, $K_{\mathbf{X},2}(\mathbf{x})$, $K_{\mathbf{X},4}(\mathbf{x})$ et $K_{\mathbf{X},5}(\mathbf{x})$ sont des constantes qui se retrouvent dans l'intégrale du numérateur et dans celle du dénominateur, alors elles s'annuleront et nous pouvons par conséquent les ignorer dans le calcul de la fonction de poids. Toutefois nous devons calculer $K_{\mathbf{X},3}(\epsilon_x, \mathbf{x})$ car elle dépend de ϵ_x , une variable d'intégration. Nous trouvons $w_{\mathbf{X},1}(\mu_{x_1}, \sigma_{x_1}^2, \epsilon_x, \mathbf{z})$ de la même façon en changeant x_2 pour x_1 et z_i pour $1 - z_i$ et nous trouvons $w_{\mathbf{Y},1}(\mu_{y_1}, \sigma_{y_1}^2, \epsilon_y, \mathbf{w})$ et $w_{\mathbf{Y},2}(\mu_{y_2}, \sigma_{y_2}^2, \epsilon_y, \mathbf{w})$ de la même façon en changeant \mathbf{x} par \mathbf{y} et \mathbf{z} par \mathbf{w} .

Nous avons tous les éléments pour estimer $p(\mathbf{x}, \mathbf{y}) = \mathbb{P}[H_0 : \mathbb{E}(X) < \mathbb{E}(Y) \mid \mathbf{x}, \mathbf{y}]$. Voici l'algorithme à suivre:

1) Choisir les paramètres des lois *a priori* de $\mu_{x_1}, \mu_{x_2}, \mu_{y_1}$ et μ_{y_2} , c'est-à-dire $(\lambda_{x_1}, \tau_{x_1}^2)$, $(\lambda_{x_2}, \tau_{x_2}^2)$, $(\lambda_{y_1}, \tau_{y_1}^2)$ et $(\lambda_{y_2}, \tau_{y_2}^2)$, les paramètres des lois *a priori* de $\sigma_{x_1}^2, \sigma_{x_2}^2, \sigma_{y_1}^2$ et $\sigma_{y_2}^2$, c'est-à-dire $(\alpha_{x_1}, \beta_{x_1}, \alpha_{x_2}, \beta_{x_2}, \alpha_{y_1}, \beta_{y_1}, \alpha_{y_2}, \beta_{y_2})$, puis les paramètres des lois *a priori* de ϵ_x et ϵ_y , c'est-à-dire c_x, d_x, c_y, d_y .

2) Déterminer $\hat{\mu}_{x_1}, \hat{\mu}_{x_2}, \hat{\mu}_{y_1}, \hat{\mu}_{y_2}, \hat{\sigma}_{x_1}^2, \hat{\sigma}_{x_2}^2, \hat{\sigma}_{y_1}^2, \hat{\sigma}_{y_2}^2, \hat{\epsilon}_x$ et $\hat{\epsilon}_y$ par l'algorithme EM.

3) Générer ϵ_x selon une loi

$$Be(c_x + m\hat{\epsilon}_x, d_x + m(1 - \hat{\epsilon}_x)) \tag{3.6.13}$$

et ϵ_y selon une loi

$$Be(c_y + n\hat{\epsilon}_y, d_y + n(1 - \hat{\epsilon}_y)). \quad (3.6.14)$$

4) Générer z_i , pour $i = 1, \dots, m$, selon une loi

$$Bin \left(1, \frac{\frac{\epsilon_x}{\hat{\sigma}_{x_2}} \exp \left[\frac{-1}{2\hat{\sigma}_{x_2}^2} (\log x_i - \hat{\mu}_{x_2})^2 \right]}{\frac{1-\epsilon_x}{\hat{\sigma}_{x_1}} \exp \left[\frac{-1}{2\hat{\sigma}_{x_1}^2} (\log x_i - \hat{\mu}_{x_1})^2 \right] + \frac{\epsilon_x}{\hat{\sigma}_{x_2}} \exp \left[\frac{-1}{2\hat{\sigma}_{x_2}^2} (\log x_i - \hat{\mu}_{x_2})^2 \right]} \right) \quad (3.6.15)$$

et w_i , pour $i = 1, \dots, n$, selon une loi

$$Bin \left(1, \frac{\frac{\epsilon_y}{\hat{\sigma}_{y_2}} \exp \left[\frac{-1}{2\hat{\sigma}_{y_2}^2} (\log y_i - \hat{\mu}_{y_2})^2 \right]}{\frac{1-\epsilon_y}{\hat{\sigma}_{y_1}} \exp \left[\frac{-1}{2\hat{\sigma}_{y_1}^2} (\log y_i - \hat{\mu}_{y_1})^2 \right] + \frac{\epsilon_y}{\hat{\sigma}_{y_2}} \exp \left[\frac{-1}{2\hat{\sigma}_{y_2}^2} (\log y_i - \hat{\mu}_{y_2})^2 \right]} \right) \quad (3.6.16)$$

5) Générer $\sigma_{x_1}^2$ selon une loi

$$IG \left(\frac{\sum(1 - z_i)}{2} + \alpha_{x_1}, \frac{1}{2} \sum (\log x_i - \hat{\mu}_{x_1})^2 (1 - z_i) + \beta_{x_1} \right), \quad (3.6.17)$$

$\sigma_{x_2}^2$ selon une loi

$$IG \left(\frac{\sum z_i}{2} + \alpha_{x_2}, \frac{1}{2} \sum (\log x_i - \hat{\mu}_{x_2})^2 z_i + \beta_{x_2} \right), \quad (3.6.18)$$

$\sigma_{y_1}^2$ selon une loi

$$IG \left(\frac{\sum(1 - w_i)}{2} + \alpha_{y_1}, \frac{1}{2} \sum (\log y_i - \hat{\mu}_{y_1})^2 (1 - w_i) + \beta_{y_1} \right) \quad (3.6.19)$$

et $\sigma_{y_2}^2$ selon une loi

$$IG \left(\frac{\sum w_i}{2} + \alpha_{y_2}, \frac{1}{2} \sum (\log y_i - \hat{\mu}_{y_2})^2 w_i + \beta_{y_2} \right). \quad (3.6.20)$$

6) Générer μ_{x_1} selon une loi

$$N \left(\frac{\bar{x}_{(1)} \tau_{x_1}^2 \sum(1 - z_i) + \lambda_{x_1} \sigma_{x_1}^2}{\tau_{x_1}^2 \sum(1 - z_i) + \sigma_{x_1}^2}, \frac{\tau_{x_1}^2 \sigma_{x_1}^2}{\tau_{x_1}^2 \sum(1 - z_i) + \sigma_{x_1}^2} \right), \quad (3.6.21)$$

μ_{x_2} selon une loi

$$N \left(\frac{\bar{x}_{(2)} \tau_{x_2}^2 \sum z_i + \lambda_{x_2} \sigma_{x_2}^2}{\tau_{x_2}^2 \sum z_i + \sigma_{x_2}^2}, \frac{\tau_{x_2}^2 \sigma_{x_2}^2}{\tau_{x_2}^2 \sum z_i + \sigma_{x_2}^2} \right), \quad (3.6.22)$$

μ_{y_1} selon une loi

$$N \left(\frac{\bar{y}_{(1)} \tau_{y_1}^2 \sum(1 - w_i) + \lambda_{y_1} \sigma_{y_1}^2}{\tau_{y_1}^2 \sum(1 - w_i) + \sigma_{y_1}^2}, \frac{\tau_{y_1}^2 \sigma_{y_1}^2}{\tau_{y_1}^2 \sum(1 - w_i) + \sigma_{y_1}^2} \right) \quad (3.6.23)$$

et μ_{y_2} selon une loi

$$N\left(\frac{\bar{y}_{(2)}\tau_{y_2}^2 \sum w_i + \lambda_{y_2}\sigma_{y_2}^2}{\tau_{y_2}^2 \sum w_i + \sigma_{y_2}^2}, \frac{\tau_{y_2}^2 \sigma_{y_2}^2}{\tau_{y_2}^2 \sum w_i + \sigma_{y_2}^2}\right). \quad (3.6.24)$$

7) Calculer les fonctions indicatrices suivantes:

$$\begin{aligned} I_{XY}(\boldsymbol{\mu}, \boldsymbol{\sigma}^2, \boldsymbol{\epsilon}) &= \begin{cases} 1 & ; \text{ si } (1 - \epsilon_x)e^{\mu_{x_1} + \sigma_{x_1}^2/2} + \epsilon_x e^{\mu_{x_2} + \sigma_{x_2}^2/2} \\ & < (1 - \epsilon_y)e^{\mu_{y_1} + \sigma_{y_1}^2/2} + \epsilon_y e^{\mu_{y_2} + \sigma_{y_2}^2/2}, \\ 0 & ; \text{ sinon,} \end{cases} \\ I_X(\boldsymbol{\mu}_x, \boldsymbol{\sigma}_x^2) &= \begin{cases} 1 & ; \text{ si } e^{\mu_{x_1} + \sigma_{x_1}^2/2} < e^{\mu_{x_2} + \sigma_{x_2}^2/2}, \\ 0 & ; \text{ sinon,} \end{cases} \\ I_Y(\boldsymbol{\mu}_y, \boldsymbol{\sigma}_y^2) &= \begin{cases} 1 & ; \text{ si } e^{\mu_{y_1} + \sigma_{y_1}^2/2} < e^{\mu_{y_2} + \sigma_{y_2}^2/2}, \\ 0 & ; \text{ sinon.} \end{cases} \end{aligned}$$

8) Calculer

$$\begin{aligned} w_{\mathbf{X},1}(\mu_{x_1}, \sigma_{x_1}^2, \epsilon_x, \mathbf{z}) &= \exp\left[\frac{1}{2\sigma_{x_1}^2}(\hat{\mu}_{x_1} - \bar{x}_{(1)})^2 \sum_{i=1}^m (1 - z_i)\right] \exp\left[-\frac{(\sum_i 1 - z_i)(\bar{x}_{(1)} - \lambda_{x_1})^2}{2(\tau_{x_1}^2 \sum_i (1 - z_i) + \sigma_{x_1}^2)}\right] \\ &\times \frac{\sigma_{x_1}}{\sqrt{\tau_{x_1}^2 \sum (1 - z_i) + \sigma_{x_1}^2}} \cdot \frac{\beta_{x_1}^{\alpha_{x_1}}}{\Gamma(\alpha_{x_1})} \cdot \frac{\Gamma(\frac{\sum(1-z_i)}{2} + \alpha_{x_1})}{\left[\frac{1}{2} \sum (\log x_i - \hat{\mu}_{x_1})^2 z_i + \beta_{x_1}\right]^{\frac{\sum(1-z_i)}{2} + \alpha_{x_1}}} \\ &\times \left(\frac{\epsilon_x^{m\hat{\epsilon}_x}}{(\hat{\sigma}_{x_1})^{\sum(1-z_i)}} \exp\left[-\frac{1}{2\hat{\sigma}_{x_1}^2}(\log x_i - \hat{\mu}_{x_1})^2 (1 - z_i)\right]\right)^{-1} \end{aligned} \quad (3.6.25)$$

et

$$\begin{aligned} w_{\mathbf{X},2}(\mu_{x_2}, \sigma_{x_2}^2, \epsilon_x, \mathbf{z}) &= \exp\left[\frac{1}{2\sigma_{x_2}^2}(\hat{\mu}_{x_2} - \bar{x}_{(2)})^2 \sum_{i=1}^m z_i\right] \exp\left[-\frac{(\sum_i z_i)(\bar{x}_{(2)} - \lambda_{x_2})^2}{2(\tau_{x_2}^2 \sum_i z_i + \sigma_{x_2}^2)}\right] \\ &\times \frac{\sigma_{x_2}}{\sqrt{\tau_{x_2}^2 \sum z_i + \sigma_{x_2}^2}} \cdot \frac{\beta_{x_2}^{\alpha_{x_2}}}{\Gamma(\alpha_{x_2})} \cdot \frac{\Gamma(\frac{\sum z_i}{2} + \alpha_{x_2})}{\left[\frac{1}{2} \sum (\log x_i - \hat{\mu}_{x_2})^2 z_i + \beta_{x_2}\right]^{\frac{\sum z_i}{2} + \alpha_{x_2}}} \\ &\times \left(\frac{\epsilon_x^{m\hat{\epsilon}_x}}{(\hat{\sigma}_{x_2})^{\sum z_i}} \exp\left[-\frac{1}{2\hat{\sigma}_{x_2}^2}(\log x_i - \hat{\mu}_{x_2})^2 z_i\right]\right)^{-1} \end{aligned} \quad (3.6.26)$$

9) Calculer

$$\begin{aligned}
& w_{Y,1}(\mu_{y_1}, \sigma_{y_1}^2, \epsilon_y, \mathbf{w}) \\
&= \exp \left[\frac{1}{2\sigma_{y_1}^2} (\hat{\mu}_{y_1} - \bar{y}_{(1)})^2 \sum_{i=1}^m (1 - w_i) \right] \exp \left[-\frac{(\sum_i (1 - w_i) (\bar{y}_{(1)} - \lambda_{y_1})^2)}{2(\tau_{y_1}^2 \sum_i (1 - w_i) + \sigma_{y_1}^2)} \right] \\
&\times \frac{\sigma_{y_1}}{\sqrt{\tau_{y_1}^2 \sum_i (1 - w_i) + \sigma_{y_1}^2}} \cdot \frac{\beta_{y_1}^{\alpha_{y_1}}}{\Gamma(\alpha_{y_1})} \cdot \frac{\Gamma(\frac{\sum_i (1 - w_i)}{2} + \alpha_{y_1})}{\left[\frac{1}{2} \sum (\log y_i - \hat{\mu}_{y_1})^2 w_i + \beta_{y_1} \right]^{\frac{\sum_i (1 - w_i)}{2} + \alpha_{y_1}}} \\
&\times \left(\frac{\epsilon_y^{n(1 - \hat{\epsilon}_y)}}{(\hat{\sigma}_{y_1})^{\sum_i (1 - w_i)}} \exp \left[-\frac{1}{2\hat{\sigma}_{y_1}^2} (\log y_i - \hat{\mu}_{y_1})^2 (1 - w_i) \right] \right)^{-1} \quad (3.6.27)
\end{aligned}$$

et

$$\begin{aligned}
& w_{Y,2}(\mu_{y_2}, \sigma_{y_2}^2, \epsilon_y, \mathbf{w}) \\
&= \exp \left[\frac{1}{2\sigma_{y_2}^2} (\hat{\mu}_{y_2} - \bar{y}_{(2)})^2 \sum_{i=1}^m w_i \right] \exp \left[-\frac{(\sum_i w_i) (\bar{y}_{(2)} - \lambda_{y_2})^2}{2(\tau_{y_2}^2 \sum_i w_i + \sigma_{y_2}^2)} \right] \\
&\times \frac{\sigma_{y_2}}{\sqrt{\tau_{y_2}^2 \sum_i w_i + \sigma_{y_2}^2}} \cdot \frac{\beta_{y_2}^{\alpha_{y_2}}}{\Gamma(\alpha_{y_2})} \cdot \frac{\Gamma(\frac{\sum_i w_i}{2} + \alpha_{y_2})}{\left[\frac{1}{2} \sum (\log y_i - \hat{\mu}_{y_2})^2 w_i + \beta_{y_2} \right]^{\frac{\sum_i w_i}{2} + \alpha_{y_2}}} \\
&\times \left(\frac{\epsilon_y^{n\hat{\epsilon}_y}}{(\hat{\sigma}_{y_2})^{\sum_i w_i}} \exp \left[-\frac{1}{2\hat{\sigma}_{y_2}^2} (\log y_i - \hat{\mu}_{y_2})^2 w_i \right] \right)^{-1}. \quad (3.6.28)
\end{aligned}$$

10) Calculer

$$\begin{aligned}
& K_{X,3}(\epsilon_x, \mathbf{x}) = \\
& \prod_{i=1}^m \left(\frac{1 - \epsilon_x}{\hat{\sigma}_{x_1}} \exp \left[\frac{-1}{2\hat{\sigma}_{x_1}^2} (\log x_i - \hat{\mu}_{x_1})^2 \right] + \frac{\epsilon_x}{\hat{\sigma}_{x_2}} \exp \left[\frac{-1}{2\hat{\sigma}_{x_2}^2} (\log x_i - \hat{\mu}_{x_2})^2 \right] \right), \quad (3.6.29)
\end{aligned}$$

$$\begin{aligned}
& K_{Y,3}(\epsilon_y, \mathbf{y}) = \\
& \prod_{i=1}^n \left(\frac{1 - \epsilon_y}{\hat{\sigma}_{y_1}} \exp \left[\frac{-1}{2\hat{\sigma}_{y_1}^2} (\log y_i - \hat{\mu}_{y_1})^2 \right] + \frac{\epsilon_y}{\hat{\sigma}_{y_2}} \exp \left[\frac{-1}{2\hat{\sigma}_{y_2}^2} (\log y_i - \hat{\mu}_{y_2})^2 \right] \right). \quad (3.6.30)
\end{aligned}$$

11) Calculer

$$\begin{aligned}
 w_j &= w_{\mathbf{X},1}(\mu_{x_1}, \sigma_{x_1}^2, \epsilon_x, \mathbf{z}) w_{\mathbf{X},2}(\mu_{x_2}, \sigma_{x_2}^2, \epsilon_x, \mathbf{z}) w_{\mathbf{Y},1}(\mu_{y_1}, \sigma_{y_1}^2, \epsilon_y, \mathbf{w}) \\
 &\quad \times w_{\mathbf{Y},2}(\mu_{y_2}, \sigma_{y_2}^2, \epsilon_y, \mathbf{w}) K_{\mathbf{X},3}(\epsilon_x, \mathbf{x}) K_{\mathbf{Y},3}(\epsilon_y, \mathbf{y}) I_X(\mu_{\mathbf{x}}, \sigma_{\mathbf{x}}^2) I_Y(\mu_{\mathbf{y}}, \sigma_{\mathbf{y}}^2), \\
 q_j &= I_{XY}(\boldsymbol{\mu}, \boldsymbol{\sigma}^2, \boldsymbol{\epsilon}).
 \end{aligned}$$

12) Les étapes 3 à 11 représentent une simulation notée j . Refaire ces étapes un grand nombre de fois, disons S , afin de générer une suite w_1, \dots, w_S et une suite q_1, \dots, q_S .

13) Estimer la probabilité *a posteriori* de l'hypothèse H_0 par la formule suivante:

$$\mathbb{P}[H_0 : \mathbb{E}(X) < \mathbb{E}(Y) \mid \mathbf{x}, \mathbf{y}] = \frac{\sum_{j=1}^S q_j w_j}{\sum_{j=1}^S w_j}.$$

3.7. TEST D'HYPOTHÈSES BILATÉRAL AVEC POIDS ET VARIANCES ALÉATOIRES

Comme dans le test unilatéral de la section 3.6, les paramètres $\sigma_{x_1}^2, \sigma_{x_2}^2, \sigma_{y_1}^2, \sigma_{y_2}^2, \epsilon_x$ et ϵ_y sont maintenant considérés comme étant des variables aléatoires en plus des paramètres $\mu_{x_1}, \mu_{x_2}, \mu_{y_1}, \mu_{y_2}, \mathbf{z}$ et \mathbf{w} qui étaient déjà considérés aléatoires dans le test bilatéral de la section 3.5. Aussi le modèle est le même et le test bilatéral sera semblable en de nombreux points au test unilatéral.

Le test d'hypothèses est $H_0 : \mathbb{E}(X) = \mathbb{E}(Y)$ contre $H_1 : \mathbb{E}(X) \neq \mathbb{E}(Y)$, où $\mathbb{E}(X)$ et $\mathbb{E}(Y)$ sont données par les équations (3.4.12) et (3.4.13) respectivement.

Notons que les lois *a priori* sur $\sigma_{x_1}^2, \sigma_{x_2}^2, \sigma_{y_1}^2, \sigma_{y_2}^2, \epsilon_x, \epsilon_y, \mathbf{z}$ et \mathbf{w} demeurent les mêmes que dans le test unilatéral de la section 3.6. Toutefois, nous ne pouvons garder les mêmes lois *a priori* continues pour $\mu_{x_1}, \mu_{x_2}, \mu_{y_1}$ et μ_{y_2} , sans quoi la probabilité *a priori* que H_0 survienne sera nulle puisque c'est une hypothèse discrète. Il faut donc poser une mesure de Dirac pour modéliser la probabilité *a priori* que H_0 survienne.

Comme pour la section 3.5, notons aussi que $H_0 : \mathbb{E}(X) = \mathbb{E}(Y)$ est équivalent à $H_0 : \mu_{y_2} = \mu_{y_2}^*$, où $\mu_{y_2}^*$ est donné par l'équation (3.5.1).

Notons p_0 la probabilité *a priori* de l'hypothèse H_0 , c'est-à-dire

$$\begin{aligned} p_0 &= \mathbb{P}[H_0 : \mathbb{E}(X) = \mathbb{E}(Y)] \\ &= \mathbb{P}[\mu_{y_2} = \mu_{y_2}^*]. \end{aligned}$$

Ainsi nous définissons les lois a priori de $\mu_{x_1}, \mu_{x_2}, \mu_{y_1}, \mu_{y_2}, \sigma_{x_1}^2, \sigma_{x_2}^2, \sigma_{y_1}^2, \sigma_{y_2}^2, \epsilon_x$ et ϵ_y de cette façon:

$$\begin{aligned} \pi(\mu_{\mathbf{x}}, \mu_{\mathbf{y}}, \sigma_{\mathbf{x}}^2, \sigma_{\mathbf{y}}^2, \epsilon_x, \epsilon_y) &= (1 - p_0)\pi_1(\mu_{\mathbf{x}}, \mu_{\mathbf{y}}, \sigma_{\mathbf{x}}^2, \sigma_{\mathbf{y}}^2, \epsilon_x, \epsilon_y) \\ &+ p_0 \pi_2(\mu_{\mathbf{x}}, \mu_{\mathbf{y}}, \sigma_{\mathbf{x}}^2, \sigma_{\mathbf{y}}^2, \epsilon_x, \epsilon_y), \end{aligned}$$

où

$$\begin{aligned} \pi_1(\mu_{\mathbf{x}}, \mu_{\mathbf{y}}, \sigma_{\mathbf{x}}^2, \sigma_{\mathbf{y}}^2, \epsilon_x, \epsilon_y) &= K_{\mathbf{X},2}(\mathbf{x})K_{\mathbf{Y},2}(\mathbf{y})\pi_{\mu_{x_1}}(\mu_{x_1})\pi_{\mu_{x_2}}(\mu_{x_2})\pi_{\mu_{y_1}}(\mu_{y_1})\pi_{\mu_{y_2}}(\mu_{y_2}) \\ &\times \pi_{\sigma_{x_1}^2}(\sigma_{x_1}^2)\pi_{\sigma_{x_2}^2}(\sigma_{x_2}^2)\pi_{\sigma_{y_1}^2}(\sigma_{y_1}^2)\pi_{\sigma_{y_2}^2}(\sigma_{y_2}^2)\pi_{\epsilon_x}(\epsilon_x)\pi_{\epsilon_y}(\epsilon_y) \\ &\times I_X(\mu_{\mathbf{x}}, \sigma_{\mathbf{x}}^2)I_Y(\mu_{\mathbf{y}}, \sigma_{\mathbf{y}}^2), \\ \pi_2(\mu_{\mathbf{x}}, \mu_{\mathbf{y}}, \sigma_{\mathbf{x}}^2, \sigma_{\mathbf{y}}^2, \epsilon_x, \epsilon_y) &= K_{\mathbf{X},\mathbf{Y}}(\mathbf{x}, \mathbf{y})\pi_{\mu_{x_1}}(\mu_{x_1})\pi_{\mu_{x_2}}(\mu_{x_2})\pi_{\mu_{y_1}}(\mu_{y_1})I_{\mu_{y_2}^*} \\ &\times \pi_{\sigma_{x_1}^2}(\sigma_{x_1}^2)\pi_{\sigma_{x_2}^2}(\sigma_{x_2}^2)\pi_{\sigma_{y_1}^2}(\sigma_{y_1}^2)\pi_{\sigma_{y_2}^2}(\sigma_{y_2}^2)\pi_{\epsilon_x}(\epsilon_x)\pi_{\epsilon_y}(\epsilon_y) \\ &\times I_X(\mu_{\mathbf{x}}, \sigma_{\mathbf{x}}^2)I_Y^*(\mu_{y_1}, \mu_{y_2}^*, \sigma_{\mathbf{y}}^2) \end{aligned}$$

où $K_{\mathbf{X},2}(\mathbf{x}), K_{\mathbf{Y},2}(\mathbf{y}), \pi_{\mu_{x_1}}(\mu_{x_1}), \pi_{\mu_{x_2}}(\mu_{x_2}), \pi_{\mu_{y_1}}(\mu_{y_1}), \pi_{\mu_{y_2}}(\mu_{y_2}), \pi_{\sigma_{x_1}^2}(\sigma_{x_1}^2), \pi_{\sigma_{x_2}^2}(\sigma_{x_2}^2), \pi_{\sigma_{y_1}^2}(\sigma_{y_1}^2), \pi_{\sigma_{y_2}^2}(\sigma_{y_2}^2), \pi_{\epsilon_x}(\epsilon_x), \pi_{\epsilon_y}(\epsilon_y), I_X(\mu_{\mathbf{x}}, \sigma_{\mathbf{x}}^2)$ et $I_Y(\mu_{\mathbf{y}}, \sigma_{\mathbf{y}}^2)$ sont définies de la même façon que dans le test unilatéral selon les équations (3.4.10) et (3.6.3) à (3.6.7). Aussi, nous avons

$$\begin{aligned} I_{\mu_{y_2}^*} &= \begin{cases} 1 & ; \text{ si } \mu_{y_2} = \mu_{y_2}^*, \\ 0 & ; \text{ sinon,} \end{cases} \\ I_Y^*(\mu_{y_1}, \mu_{y_2}^*, \sigma_{\mathbf{y}}^2) &= \begin{cases} 1 & ; \text{ si } e^{\mu_{y_1} + \sigma_{y_1}^2/2} < e^{\mu_{y_2}^* + \sigma_{y_2}^2/2}, \\ 0 & ; \text{ sinon,} \end{cases} \\ K_{\mathbf{X},\mathbf{Y}}(\mathbf{x}, \mathbf{y})^{-1} &= \int_{\epsilon} \int_{\sigma^2} \int_{\mu_{y_1}} \int_{\mu_{\mathbf{x}}} \pi_{\mu_{x_1}}(\mu_{x_1})\pi_{\mu_{x_2}}(\mu_{x_2})\pi_{\mu_{y_1}}(\mu_{y_1})\pi_{\sigma_{x_1}^2}(\sigma_{x_1}^2)\pi_{\sigma_{x_2}^2}(\sigma_{x_2}^2) \\ &\times \pi_{\sigma_{y_1}^2}(\sigma_{y_1}^2)\pi_{\sigma_{y_2}^2}(\sigma_{y_2}^2)\pi_{\epsilon_x}(\epsilon_x)\pi_{\epsilon_y}(\epsilon_y)I_X(\mu_{\mathbf{x}}, \sigma_{\mathbf{x}}^2)I_Y^*(\mu_{y_1}, \mu_{y_2}^*, \sigma_{\mathbf{y}}^2) \\ &\times d\mu_{x_1} d\mu_{x_2} d\mu_{y_1} d\sigma_{x_1}^2 d\sigma_{x_2}^2 d\sigma_{y_1}^2 d\sigma_{y_2}^2 d\epsilon_x d\epsilon_y \end{aligned}$$

où μ correspond à $(\mu_{\mathbf{x}}, \mu_{\mathbf{y}})$, σ^2 correspond à $(\sigma_{\mathbf{x}}^2, \sigma_{\mathbf{y}}^2)$ et ϵ correspond à (ϵ_x, ϵ_y) .

Nous pouvons calculer $K_{X,2}(\mathbf{x})$, $K_{Y,2}(\mathbf{y})$ et $K_{X,Y}(\mathbf{x}, \mathbf{y})$ par la méthode de Monte Carlo. Contrairement au test unilatéral, il faudra les calculer puisque ces constantes ne s'annuleront pas par un rapport d'intégrales.

Comme nous avons observé $\mathbf{x} = (x_1, \dots, x_m)^t$ et $\mathbf{y} = (y_1, \dots, y_n)^t$, nous sommes intéressés à calculer la probabilité *a posteriori* que l'hypothèse H_0 soit vraie, c'est-à-dire:

$$\begin{aligned}
p(\mathbf{x}, \mathbf{y}) &= \mathbb{P}[H_0 : \mathbb{E}(X) = \mathbb{E}(Y) \mid \mathbf{x}, \mathbf{y}] \\
&= \mathbb{P}[\mu_{y_2} = \mu_{y_2}^* \mid \mathbf{x}, \mathbf{y}] \\
&= \mathbb{E}^{\pi(\mu_{\mathbf{x}}, \mu_{\mathbf{y}}, \sigma_{\mathbf{x}}^2, \sigma_{\mathbf{y}}^2, \epsilon_x, \epsilon_y, \mathbf{z}, \mathbf{w} \mid \mathbf{x}, \mathbf{y})} (I_{\mu_{y_2}^*}) \\
&= \sum_{\mathbf{z}} \sum_{\mathbf{w}} \int_{\epsilon} \int_{\sigma^2} \int_{\mu} I_{\mu_{y_2}^*} \pi(\mu_{\mathbf{x}}, \mu_{\mathbf{y}}, \sigma_{\mathbf{x}}^2, \sigma_{\mathbf{y}}^2, \epsilon_x, \epsilon_y, \mathbf{z}, \mathbf{w} \mid \mathbf{x}, \mathbf{y}) d\mu_{\mathbf{x}} d\mu_{\mathbf{y}} \\
&\quad \times d\sigma_{\mathbf{x}}^2 d\sigma_{\mathbf{y}}^2 d\epsilon_x d\epsilon_y \\
&= \frac{\sum_{\mathbf{z}} \sum_{\mathbf{w}} \int_{\epsilon} \int_{\sigma^2} \int_{\mu} I_{\mu_{y_2}^*} G d\mu_{\mathbf{x}} d\mu_{\mathbf{y}} d\sigma_{\mathbf{x}}^2 d\sigma_{\mathbf{y}}^2 d\epsilon_x d\epsilon_y}{\sum_{\mathbf{z}} \sum_{\mathbf{w}} \int_{\epsilon} \int_{\sigma^2} \int_{\mu} G d\mu_{\mathbf{x}} d\mu_{\mathbf{y}} d\sigma_{\mathbf{x}}^2 d\sigma_{\mathbf{y}}^2 d\epsilon_x d\epsilon_y},
\end{aligned}$$

où

$$G = f_{X,Y}(\mathbf{x}, \mathbf{y} \mid \mu_{\mathbf{x}}, \mu_{\mathbf{y}}, \sigma_{\mathbf{x}}^2, \sigma_{\mathbf{y}}^2, \epsilon_x, \epsilon_y, \mathbf{z}, \mathbf{w}) \pi(\mu_{\mathbf{x}}, \mu_{\mathbf{y}}, \sigma_{\mathbf{x}}^2, \sigma_{\mathbf{y}}^2, \epsilon_x, \epsilon_y) \pi_{Z,W}(\mathbf{z}, \mathbf{w}).$$

En remplaçant $\pi(\mu_{\mathbf{x}}, \mu_{\mathbf{y}}, \sigma_{\mathbf{x}}^2, \sigma_{\mathbf{y}}^2, \epsilon_x, \epsilon_y)$ par

$$(1 - p_0) \pi_1(\mu_{\mathbf{x}}, \mu_{\mathbf{y}}, \sigma_{\mathbf{x}}^2, \sigma_{\mathbf{y}}^2, \epsilon_x, \epsilon_y) + p_0 \pi_2(\mu_{\mathbf{x}}, \mu_{\mathbf{y}}, \sigma_{\mathbf{x}}^2, \sigma_{\mathbf{y}}^2, \epsilon_x, \epsilon_y),$$

nous trouvons:

$$p(\mathbf{x}, \mathbf{y}) = \frac{p_0 J_2(\mathbf{x}, \mathbf{y})}{(1 - p_0) J_1(\mathbf{x}, \mathbf{y}) + p_0 J_2(\mathbf{x}, \mathbf{y})}$$

où

$$\begin{aligned}
J_1(\mathbf{x}, \mathbf{y}) &= \sum_{\mathbf{z}} \sum_{\mathbf{w}} \int_{\epsilon} \int_{\sigma^2} \int_{\mu} f_{X,Y}(\mathbf{x}, \mathbf{y} \mid \mu, \sigma^2, \epsilon, \mathbf{z}, \mathbf{w}) \pi_1(\mu_{\mathbf{x}}, \mu_{\mathbf{y}}, \sigma_{\mathbf{x}}^2, \sigma_{\mathbf{y}}^2, \epsilon_x, \epsilon_y) \\
&\quad \times \pi_Z(\mathbf{z}) \pi_W(\mathbf{w}) d\mu d\sigma^2 d\epsilon, \\
J_2(\mathbf{x}, \mathbf{y}) &= \sum_{\mathbf{z}} \sum_{\mathbf{w}} \int_{\epsilon} \int_{\sigma^2} \int_{\mu_{y_1}} \int_{\mu_{\mathbf{x}}} f_{X,Y}(\mathbf{x}, \mathbf{y} \mid \mu_{\mathbf{x}}, \mu_{y_1}, \mu_{y_2}^*, \sigma^2, \epsilon, \mathbf{z}, \mathbf{w}) \\
&\quad \times \pi_2(\mu_{\mathbf{x}}, \mu_{\mathbf{y}}, \sigma_{\mathbf{x}}^2, \sigma_{\mathbf{y}}^2, \epsilon_x, \epsilon_y) \pi_Z(\mathbf{z}) \pi_W(\mathbf{w}) d\mu_{\mathbf{x}} d\mu_{y_1} d\sigma^2 d\epsilon.
\end{aligned}$$

Alors en posant

$$R = \frac{J_2(\mathbf{x}, \mathbf{y})}{J_1(\mathbf{x}, \mathbf{y})},$$

nous obtenons

$$p(\mathbf{x}, \mathbf{y}) = \left(1 + \frac{1 - p_0}{p_0 R}\right)^{-1}.$$

Nous solutionnerons ce rapport d'intégrales, $R = \frac{J_2(\mathbf{x}, \mathbf{y})}{J_1(\mathbf{x}, \mathbf{y})}$, de façon similaire au test unilatéral de la section 3.6 en utilisant la méthode de Monte Carlo avec fonctions d'importance. Nous générons les variables $\mu_{x_1}, \mu_{x_2}, \mu_{y_1}, \mu_{y_2}, \sigma_{x_1}^2, \sigma_{x_2}^2, \sigma_{y_1}^2, \sigma_{y_2}^2, \epsilon_x, \epsilon_y, \mathbf{z}$ et \mathbf{w} par les mêmes fonctions d'importance que le test unilatéral, soient par les lois normales, inverse-gamma, bêta et binômiales que nous avons définies par les équations (3.6.13) à (3.6.24). Les fonctions de poids de l'intégrale $J_1(\mathbf{x}, \mathbf{y})$ sont identiques à celles du test unilatéral. Par contre, les fonctions de poids de $J_2(\mathbf{x}, \mathbf{y})$ changeront un peu.

Ainsi pour $J_1(\mathbf{x}, \mathbf{y})$, la fonction de poids est donnée par:

$$\begin{aligned} & w_{J_1(\mathbf{x}, \mathbf{y})}(\mu_{\mathbf{x}}, \mu_{\mathbf{y}}, \sigma_{\mathbf{x}}^2, \sigma_{\mathbf{y}}^2, \epsilon_x, \epsilon_y, \mathbf{z}, \mathbf{w}) \\ &= w_{\mathbf{X},1}(\mu_{x_1}, \sigma_{x_1}^2, \epsilon_x, \mathbf{z}) w_{\mathbf{X},2}(\mu_{x_2}, \sigma_{x_2}^2, \epsilon_x, \mathbf{z}) K_{\mathbf{X},1}(\mathbf{x}) K_{\mathbf{X},2}(\mathbf{x}) K_{\mathbf{X},3}(\epsilon_x, \mathbf{x}) \\ &\times K_{\mathbf{X},4}(\mathbf{x}) K_{\mathbf{X},5}(\mathbf{x})^{-1} I_{\mathbf{X}}(\mu_{\mathbf{x}}, \sigma_{\mathbf{x}}^2) w_{\mathbf{Y},1}(\mu_{y_1}, \sigma_{y_1}^2, \epsilon_y, \mathbf{w}) w_{\mathbf{Y},2}(\mu_{y_2}, \sigma_{y_2}^2, \epsilon_y, \mathbf{w}) K_{\mathbf{Y},1}(\mathbf{y}) \\ &\times K_{\mathbf{Y},2}(\mathbf{y}) K_{\mathbf{Y},3}(\epsilon_y, \mathbf{y}) K_{\mathbf{Y},4}(\mathbf{y}) K_{\mathbf{Y},5}(\mathbf{y})^{-1} I_{\mathbf{Y}}(\mu_{\mathbf{y}}, \sigma_{\mathbf{y}}^2), \end{aligned}$$

où $w_{\mathbf{X},1}(\mu_{x_1}, \sigma_{x_1}^2, \epsilon_x, \mathbf{z})$, $w_{\mathbf{X},2}(\mu_{x_2}, \sigma_{x_2}^2, \epsilon_x, \mathbf{z})$, $w_{\mathbf{Y},1}(\mu_{y_1}, \sigma_{y_1}^2, \epsilon_y, \mathbf{w})$, $w_{\mathbf{Y},2}(\mu_{y_2}, \sigma_{y_2}^2, \epsilon_y, \mathbf{w})$, $K_{\mathbf{X},3}(\epsilon_x, \mathbf{x})$ et $K_{\mathbf{Y},3}(\epsilon_y, \mathbf{y})$ sont donnés par les équations (3.6.25) à (3.6.30), tandis que pour $J_2(\mathbf{x}, \mathbf{y})$, nous pouvons démontrer (détails à l'annexe 1) que la fonction de poids est donnée par

$$\begin{aligned} & w_{J_2(\mathbf{x}, \mathbf{y})}(\mu_{\mathbf{x}}, \mu_{y_1}, \sigma_{\mathbf{x}}^2, \sigma_{y_1}^2, \epsilon_x, \epsilon_y, \mathbf{z}, \mathbf{w}) \\ &= w_{\mathbf{X},1}(\mu_{x_1}, \sigma_{x_1}^2, \epsilon_x, \mathbf{z}) w_{\mathbf{X},2}(\mu_{x_2}, \sigma_{x_2}^2, \epsilon_x, \mathbf{z}) w_{\mathbf{Y},1}(\mu_{y_1}, \sigma_{y_1}^2, \epsilon_y, \mathbf{w}) w'_{\mathbf{Y},2}(\mu_{y_2}^*, \sigma_{y_2}^2, \epsilon_y, \mathbf{w}) \\ &\times K_{\mathbf{X},1}(\mathbf{x}) K_{\mathbf{X},3}(\epsilon_x, \mathbf{x}) K_{\mathbf{X},4}(\mathbf{x}) K_{\mathbf{X},5}(\mathbf{x})^{-1} K_{\mathbf{Y},1}(\mathbf{y}) K_{\mathbf{Y},3}(\epsilon_y, \mathbf{y}) K_{\mathbf{Y},4}(\mathbf{y}) K_{\mathbf{Y},5}(\mathbf{y})^{-1} \\ &\times K_{\mathbf{X},\mathbf{Y}}(\mathbf{x}, \mathbf{y}) I_{\mathbf{X}}(\mu_{\mathbf{x}}, \sigma_{\mathbf{x}}^2) I_{\mathbf{Y}}^*(\mu_{y_1}, \mu_{y_2}^*, \sigma_{y_1}^2) \end{aligned}$$

où

$$\begin{aligned}
w'_{\mathbf{Y},2}(\mu_{y_2}^*, \sigma_{y_2}^2, \epsilon_y, \mathbf{w}) &= \exp\left(-\frac{1}{2\sigma_{y_2}^2} \sum w_i [(\log y_i - \mu_{y_2}^*)^2 - (\log y_i - \hat{\mu}_{y_2})^2]\right) \\
&\times \frac{\beta_{y_2}^{\alpha_{y_2}}}{\Gamma(\alpha_{y_2})} \cdot \frac{\Gamma(\frac{\sum w_i}{2} + \alpha_{y_2})}{\left[\frac{1}{2} \sum (\log y_i - \hat{\mu}_{y_2})^2 w_i + \beta_{y_2}\right]^{\frac{\sum w_i}{2} + \alpha_{y_2}}} \\
&\times \left(\frac{\epsilon_y^{n\hat{\epsilon}_y}}{\hat{\sigma}_{y_2}^{\sum w_i}} \exp\left[-\frac{1}{2\hat{\sigma}_{y_2}^2} \sum (\log y_i - \hat{\mu}_{y_2})^2 w_i\right]\right)^{-1}.
\end{aligned}$$

Par la méthode de Monte Carlo avec fonctions d'importance, nous estimons R par

$$R = \frac{J_2(\mathbf{x}, \mathbf{y})}{J_1(\mathbf{x}, \mathbf{y})} \approx \frac{\sum_{j=1}^S w_{J_2(\mathbf{x}, \mathbf{y})(j)}(\mu_{\mathbf{x}}, \mu_{y_1}, \sigma_{\mathbf{x}}^2, \sigma_{y_1}^2, \epsilon_x, \epsilon_y, \mathbf{z}, \mathbf{w})}{\sum_{j=1}^S w_{J_1(\mathbf{x}, \mathbf{y})(j)}(\mu_{\mathbf{x}}, \mu_{\mathbf{y}}, \sigma_{\mathbf{x}}^2, \sigma_{y_2}^2, \epsilon_x, \epsilon_y, \mathbf{z}, \mathbf{w})}.$$

Les constantes $K_{\mathbf{X},1}(\mathbf{x})$, $K_{\mathbf{X},4}(\mathbf{x})$, $K_{\mathbf{X},5}(\mathbf{x})$, $K_{\mathbf{Y},1}(\mathbf{y})$, $K_{\mathbf{Y},4}(\mathbf{y})$ et $K_{\mathbf{Y},5}(\mathbf{y})$ se retrouvent au numérateur et au dénominateur du rapport R et nous pouvons par conséquent les ignorer dans le calcul de la fonction de poids. Nous avons tous les éléments pour estimer $p(\mathbf{x}, \mathbf{y}) = \mathbb{P}[H_0 : \mathbb{E}(X) = \mathbb{E}(Y) \mid \mathbf{x}, \mathbf{y}]$. Voici l'algorithme à suivre:

1) Choisir les paramètres des lois *a priori* de $\mu_{x_1}, \mu_{x_2}, \mu_{y_1}$ et μ_{y_2} , c'est-à-dire $(\lambda_{x_1}, \tau_{x_1}^2)$, $(\lambda_{x_2}, \tau_{x_2}^2)$, $(\lambda_{y_1}, \tau_{y_1}^2)$ et $(\lambda_{y_2}, \tau_{y_2}^2)$, les paramètres des lois *a priori* de $\sigma_{x_1}^2, \sigma_{x_2}^2, \sigma_{y_1}^2$ et $\sigma_{y_2}^2$, c'est-à-dire $(\alpha_{x_1}, \beta_{x_1}, \alpha_{x_2}, \beta_{x_2}, \alpha_{y_1}, \beta_{y_1}, \alpha_{y_2}, \beta_{y_2})$ puis les paramètres des lois *a priori* de ϵ_x et ϵ_y , c'est-à-dire c_x, d_x, c_y, d_y .

2) Déterminer $\hat{\mu}_{x_1}, \hat{\mu}_{x_2}, \hat{\mu}_{y_1}, \hat{\mu}_{y_2}, \hat{\sigma}_{x_1}^2, \hat{\sigma}_{x_2}^2, \hat{\sigma}_{y_1}^2, \hat{\sigma}_{y_2}^2, \hat{\epsilon}_x$ et $\hat{\epsilon}_y$ par l'algorithme EM.

3) Déterminer $p_0 = \mathbb{P}[H_0 : \mathbb{E}(X) = \mathbb{E}(Y)]$ la probabilité *a priori* de l'hypothèse H_0 .

4) Calculer par la méthode de Monte Carlo les constantes suivantes:

$$\begin{aligned} K_{\mathbf{X},\mathbf{Y}}(\mathbf{x}, \mathbf{y})^{-1} &= \int_{\epsilon} \int_{\sigma^2} \int_{\mu_{y_1}} \int_{\mu_{\mathbf{x}}} \pi_{\mu_{x_1}}(\mu_{x_1}) \pi_{\mu_{x_2}}(\mu_{x_2}) \pi_{\mu_{y_1}}(\mu_{y_1}) \pi_{\sigma_{x_1}^2}(\sigma_{x_1}^2) \pi_{\sigma_{x_2}^2}(\sigma_{x_2}^2) \\ &\times \pi_{\sigma_{y_1}^2}(\sigma_{y_1}^2) \pi_{\sigma_{y_2}^2}(\sigma_{y_2}^2) \pi_{\epsilon_x}(\epsilon_x) \pi_{\epsilon_y}(\epsilon_y) I_X(\mu_{\mathbf{x}}, \sigma_{\mathbf{x}}^2) I_Y^*(\mu_{y_1}, \mu_{y_2}^*, \sigma_{\mathbf{y}}^2) \\ &\times d\mu_{x_1} d\mu_{x_2} d\mu_{y_1} d\sigma_{x_1}^2 d\sigma_{x_2}^2 d\sigma_{y_1}^2 d\sigma_{y_2}^2 d\epsilon_x d\epsilon_y, \end{aligned}$$

$$\begin{aligned} K_{\mathbf{X},2}(\mathbf{x})^{-1} &= \int_0^\infty \int_0^\infty \int_{-\infty}^\infty \int_{-\infty}^\infty \pi_{\mu_{x_1}}(\mu_{x_1}) \pi_{\mu_{x_2}}(\mu_{x_2}) \pi_{\sigma_{x_1}^2}(\sigma_{x_1}^2) \pi_{\sigma_{x_2}^2}(\sigma_{x_2}^2) I_X(\mu_{\mathbf{x}}, \sigma_{\mathbf{x}}^2) \\ &\times d\mu_{x_1} d\mu_{x_2} d\sigma_{x_1}^2 d\sigma_{x_2}^2, \end{aligned}$$

$$\begin{aligned} K_{\mathbf{Y},2}(\mathbf{y})^{-1} &= \int_0^\infty \int_0^\infty \int_{-\infty}^\infty \int_{-\infty}^\infty \pi_{\mu_{y_1}}(\mu_{y_1}) \pi_{\mu_{y_2}}(\mu_{y_2}) \pi_{\sigma_{y_1}^2}(\sigma_{y_1}^2) \pi_{\sigma_{y_2}^2}(\sigma_{y_2}^2) I_Y(\mu_{\mathbf{y}}, \sigma_{\mathbf{y}}^2) \\ &\times d\mu_{y_1} d\mu_{y_2} d\sigma_{y_1}^2 d\sigma_{y_2}^2. \end{aligned}$$

5) Refaire les étapes 3 à 6 du test unilatéral de la section 3.6, soit générer $\epsilon_x, \epsilon_y, z_i$, pour $i = 1, \dots, m$, w_i pour $i = 1, \dots, n$, $\sigma_{x_1}^2, \sigma_{x_2}^2, \sigma_{y_1}^2, \sigma_{y_2}^2, \mu_{x_1}, \mu_{x_2}, \mu_{y_1}$ et μ_{y_2} selon les équations (3.6.13) à (3.6.24).

6) Calculer

$$\mu_{y_2}^* = \log \left[\frac{(1 - \epsilon_x) e^{\mu_{x_1} + \sigma_{x_1}^2/2} + \epsilon_x e^{\mu_{x_2} + \sigma_{x_2}^2/2} - (1 - \epsilon_y) e^{\mu_{y_1} + \sigma_{y_1}^2/2}}{\epsilon_y} \right] - \sigma_{y_2}^2/2.$$

7) Calculer les fonctions indicatrices suivantes:

$$\begin{aligned} I_X(\mu_{\mathbf{x}}, \sigma_{\mathbf{x}}^2) &= \begin{cases} 1 & ; \text{ si } e^{\mu_{x_1} + \sigma_{x_1}^2/2} < e^{\mu_{x_2} + \sigma_{x_2}^2/2}, \\ 0 & ; \text{ sinon,} \end{cases} \\ I_Y(\mu_{\mathbf{y}}, \sigma_{\mathbf{y}}^2) &= \begin{cases} 1 & ; \text{ si } e^{\mu_{y_1} + \sigma_{y_1}^2/2} < e^{\mu_{y_2} + \sigma_{y_2}^2/2}, \\ 0 & ; \text{ sinon,} \end{cases} \\ I_Y^*(\mu_{y_1}, \mu_{y_2}^*, \sigma_{\mathbf{y}}^2) &= \begin{cases} 1 & ; \text{ si } e^{\mu_{y_1} + \sigma_{y_1}^2/2} < e^{\mu_{y_2}^* + \sigma_{y_2}^2/2}, \\ 0 & ; \text{ sinon.} \end{cases} \end{aligned}$$

8) Refaire les étapes 8 à 10 du test unilatéral de la section 3.6, soit calculer les fonctions de poids $w_{\mathbf{X},1}(\mu_{x_1}, \sigma_{x_1}^2, \epsilon_x, \mathbf{z})$, $w_{\mathbf{X},2}(\mu_{x_2}, \sigma_{x_2}^2, \epsilon_x, \mathbf{z})$, $w_{\mathbf{Y},1}(\mu_{y_1}, \sigma_{y_1}^2, \epsilon_y, \mathbf{w})$, $w_{\mathbf{Y},2}(\mu_{y_2}, \sigma_{y_2}^2, \epsilon_y, \mathbf{w})$ et les constantes $K_{\mathbf{X},3}(\epsilon_x, \mathbf{x})$ et $K_{\mathbf{Y},3}(\epsilon_y, \mathbf{y})$ selon les équations (3.6.25) à (3.6.30).

9) Calculer

$$\begin{aligned}
& w'_{\mathbf{Y},2}(\mu_{y_2}^*, \sigma_{y_2}^2, \epsilon_y, \mathbf{w}) \\
&= \exp\left(-\frac{1}{2\sigma_{y_2}^2} \sum w_i [(\log y_i - \mu_{y_2}^*)^2 - (\log y_i - \hat{\mu}_{y_2})^2]\right) \\
&\times \frac{\beta_{y_2}^{\alpha_{y_2}}}{\Gamma(\alpha_{y_2})} \cdot \frac{\Gamma(\frac{\sum w_i}{2} + \alpha_{y_2})}{\left[\frac{1}{2} \sum (\log y_i - \hat{\mu}_{y_2})^2 w_i + \beta_{y_2}\right]^{\frac{\sum w_i}{2} + \alpha_{y_2}}} \\
&\times \left(\frac{\epsilon_y^{n\hat{\epsilon}_y}}{\hat{\sigma}_{y_2}^{\sum w_i}} \exp\left[-\frac{1}{2\hat{\sigma}_{y_2}^2} \sum (\log y_i - \hat{\mu}_{y_2})^2 w_i\right]\right)^{-1}.
\end{aligned}$$

10) Calculer

$$\begin{aligned}
w'_j &= K_{\mathbf{X},2}(\mathbf{x}) K_{\mathbf{Y},2}(\mathbf{y}) w_{\mathbf{X},1}(\mu_{x_1}, \sigma_{x_1}^2, \epsilon_x, \mathbf{z}) w_{\mathbf{X},2}(\mu_{x_2}, \sigma_{x_2}^2, \epsilon_x, \mathbf{z}) w_{\mathbf{Y},1}(\mu_{y_1}, \sigma_{y_1}^2, \epsilon_y, \mathbf{w}) \\
&\times w_{\mathbf{Y},2}(\mu_{y_2}, \sigma_{y_2}^2, \epsilon_y, \mathbf{w}) K_{\mathbf{X},3}(\epsilon_x, \mathbf{x}) K_{\mathbf{Y},3}(\epsilon_y, \mathbf{y}) I_X(\mu_{\mathbf{x}}, \sigma_{\mathbf{x}}^2) I_Y(\mu_{\mathbf{y}}, \sigma_{\mathbf{y}}^2), \\
w''_j &= K_{\mathbf{X},\mathbf{Y}}(\mathbf{x}, \mathbf{y}) w_{\mathbf{X},1}(\mu_{x_1}, \sigma_{x_1}^2, \epsilon_x, \mathbf{z}) w_{\mathbf{X},2}(\mu_{x_2}, \sigma_{x_2}^2, \epsilon_x, \mathbf{z}) w_{\mathbf{Y},1}(\mu_{y_1}, \sigma_{y_1}^2, \epsilon_y, \mathbf{w}) \\
&\times w'_{\mathbf{Y},2}(\mu_{y_2}^*, \sigma_{y_2}^2, \epsilon_y, \mathbf{w}) K_{\mathbf{X},3}(\epsilon_x, \mathbf{x}) K_{\mathbf{Y},3}(\epsilon_y, \mathbf{y}) I_X(\mu_{\mathbf{x}}, \sigma_{\mathbf{x}}^2) I_Y^*(\mu_{y_1}, \mu_{y_2}^*, \sigma_{\mathbf{y}}^2).
\end{aligned}$$

11) Les étapes 5 à 10 représentent une simulation notée j . Refaire ces étapes un grand nombre de fois, disons S , afin de générer une suite w'_1, \dots, w'_S et une suite w''_1, \dots, w''_S .

12) Estimer la probabilité par les formules suivantes:

$$R = \frac{\sum_{j=1}^S w''_j}{\sum_{j=1}^S w'_j}$$

et

$$p(\mathbf{x}, \mathbf{y}) = \mathbb{P}[H_0 : \mathbb{E}(X) = \mathbb{E}(Y) \mid \mathbf{x}, \mathbf{y}] = \left(1 + \frac{1 - p_0}{p_0 R}\right)^{-1}.$$

3.8. APPLICATION DES TESTS D'HYPOTHÈSES UNILATÉRAL ET BILATÉRAL SUR LES DONNÉES DE L'ÉTUDE

Nous voulons dans cette section appliquer les algorithmes présentés dans les sections précédentes sur les données de l'étude. Afin de simplifier la présentation, nous appliquerons les algorithmes des tests des sections 3.4 et 3.5 seulement, soit

les tests unilatéral et bilatéral en considérant μ_{x_1} , μ_{x_2} , μ_{y_1} , μ_{y_2} , \mathbf{z} et \mathbf{w} comme étant des variables aléatoires, mais en considérant $\sigma_{x_1}^2$, $\sigma_{x_2}^2$, $\sigma_{y_1}^2$, $\sigma_{y_2}^2$, ϵ_x et ϵ_y comme connues.

Selon l'étape 1 des deux algorithmes, nous devons choisir les paramètres de la loi *a priori* de μ_{x_1} , μ_{x_2} , μ_{y_1} et μ_{y_2} , soit les paramètres d'une loi normale. Le choix doit refléter l'information *a priori* que nous possédons. Dans cette étude, nous pensons *a priori* que le traitement 1 est moins coûteux que le traitement 2 si le traitement est un succès car son coût d'acquisition est moins élevé. C'est pourquoi nous choisirons la moyenne de μ_{x_1} inférieure à celle de μ_{y_1} . Si le traitement est un échec et s'il y a des effets secondaires, nous ne sommes pas en mesure de savoir *a priori* s'il y a une différence de coûts entre les deux traitements. C'est pourquoi nous choisirons la moyenne de μ_{x_2} égale à celle de μ_{y_2} . Le choix est donc le suivant:

$$\lambda_{x_1} = 5,40, \lambda_{x_2} = 6,30, \tau_{x_1}^2 = 0,10, \tau_{x_2}^2 = 0,15$$

et

$$\lambda_{y_1} = 5,65, \lambda_{y_2} = 6,30, \tau_{y_1}^2 = 0,10, \tau_{y_2}^2 = 0,15.$$

Notons que les lois *a priori* pour les deux traitements sont quand même assez semblables. Il est effectivement fréquent en pharmacoeconomie de placer les deux traitements sur un pied d'égalité avant d'avoir effectué l'étude.

L'étape 2 consiste à estimer les paramètres du modèle par l'algorithme EM. Les résultats de ces estimés ont déjà été présentés aux tableaux 2.1 et 2.2. Pour les étapes subséquentes, nous appliquons les algorithmes tels que décrits aux sections 3.4 et 3.5. Nous trouvons à l'annexe 2 le programme informatique écrit en XLispStat qui effectue cet algorithme. Pour le test bilatéral, nous devons fixer $p_0 = \mathbb{P}[H_0 : \mathbb{E}(X) = \mathbb{E}(Y)]$, soit la probabilité *a priori* que l'hypothèse H_0 survienne. En général, nous le fixons à 1/2 pour ne pas favoriser ni H_0 ni H_1 . Nous appliquons ces deux algorithmes avec 100 000 simulations de Monte Carlo

et nous trouvons ces résultats:

$$\begin{aligned}\mathbb{P}[H_0 : \mathbb{E}(X) < \mathbb{E}(Y) \mid \mathbf{x}, \mathbf{y}] &= 0,0044, \\ \mathbb{P}[H_0 : \mathbb{E}(X) = \mathbb{E}(Y) \mid \mathbf{x}, \mathbf{y}] &= 0,112.\end{aligned}$$

Si nous choisissons les paramètres de la fonction de coûts présentée à la section 3.1, tel que $a_0 = 19$ et $a_1 = 1$, c'est-à-dire que nous pénalisons beaucoup plus l'erreur de type I, alors nous rejettons l'hypothèse H_0 si $\mathbb{P}(H_0 \mid \mathbf{x}, \mathbf{y}) < 0,05$, soit l'analogue du test classique où on rejette H_0 si la valeur- p est plus petite que 0,05.

Nous nous rappellerons que la valeur- p de la méthode de rééchantillonnage paramétrique donnait 0,114 pour le test bilatéral et 0,005 pour le test unilatéral. Ces résultats sont très semblables aux probabilités bayésiennes. Un des avantages de l'approche bayésienne est que le résultat s'interprète facilement, puisque c'est la probabilité *a posteriori* que l'hypothèse H_0 survienne. Dans le cas des tests classiques, la valeur- p est plus difficile à interpréter. De plus, même si les lois *a priori* ne sont pas toujours faciles à choisir, elles ont l'avantage d'être explicites dans l'approche bayésienne.

3.9. PROBABILITÉ A POSTERIORI QUE LE TRAITEMENT 1 SOIT MOINS COÛTEUX QUE LE TRAITEMENT 2

Nous sommes intéressés à calculer

$$\begin{aligned}\mathbb{P}[X_{m+1} < Y_{n+1} \mid \mathbf{x}, \mathbf{y}] &= \int_{x_{m+1}} \int_{y_{n+1}} m_{X,Y}(x_{m+1}, y_{n+1} \mid \mathbf{x}, \mathbf{y}) \mathbb{I}(x_{m+1} < y_{n+1}) dy_{n+1} dx_{m+1} \\ &= \mathbb{E}^{m_{X,Y}(x_{m+1}, y_{n+1} \mid \mathbf{x}, \mathbf{y})} [\mathbb{I}(x_{m+1} < y_{n+1})]\end{aligned}$$

où

$$\begin{aligned}m_{X,Y}(x_{m+1}, y_{n+1} \mid \mathbf{x}, \mathbf{y}) &= m_X(x_{m+1} \mid \mathbf{x}) m_Y(y_{n+1} \mid \mathbf{y}), \\ \mathbb{I}(x_{m+1} < y_{n+1}) &= \begin{cases} 1 & ; \text{ si } x_{m+1} < y_{n+1}, \\ 0 & ; \text{ sinon.} \end{cases}\end{aligned}$$

Développons $m_X(x_{m+1} | \mathbf{x})$:

$$\begin{aligned}
m_X(x_{m+1} | \mathbf{x}) &= \frac{m_{\mathbf{X}}(x_1, \dots, x_m, x_{m+1})}{m_{\mathbf{X}}(x_1, \dots, x_m)} \\
&= \frac{\int_{\psi_x} f_X(x_{m+1} | \psi_x) \prod_{i=1}^m f_X(x_i | \psi_x) \pi_{\psi_x}(\psi_x) d\psi_x}{\int_{\psi_x} \prod_{i=1}^m f_X(x_i | \psi_x) \pi_{\psi_x}(\psi_x) d\psi_x} \\
&= \frac{\int_{\psi_x} f_X(x_{m+1} | \psi_x) m_{\mathbf{X}}(x_1, \dots, x_m) \pi_{\psi_x}(\psi_x | x_1, \dots, x_m) d\psi_x}{m_{\mathbf{X}}(x_1, \dots, x_m)} \\
&= \int_{\psi_x} f_X(x_{m+1} | \psi_x) \pi_{\psi_x}(\psi_x | \mathbf{x}) d\psi_x \\
&= \mathbb{E}^{\pi_{\psi_x}(\psi_x | \mathbf{x})} [f_X(x_{m+1} | \psi_x)].
\end{aligned}$$

Nous obtenons donc

$$\begin{aligned}
&\mathbb{P}[X_{m+1} < Y_{n+1} | \mathbf{x}, \mathbf{y}] \\
&= \int_{x_{m+1}} \int_{y_{n+1}} \int_{\psi_x} \int_{\psi_y} f_{X,Y}(x_{m+1}, y_{n+1} | \psi_x, \psi_y) \\
&\times \pi_{\psi_x, \psi_y}(\psi_x, \psi_y | \mathbf{x}, \mathbf{y}) \mathbb{I}(x_{m+1} < y_{n+1}) d\psi_y d\psi_x dy_{n+1} dx_{m+1} \\
&= \mathbb{E}^{\pi_{\psi_x, \psi_y}(\psi_x, \psi_y | \mathbf{x}, \mathbf{y})} \mathbb{E}^{f_{X,Y}(x_{m+1}, y_{n+1} | \psi_x, \psi_y)} [\mathbb{I}(x_{m+1} < y_{n+1})].
\end{aligned}$$

Dans l'étude considérée, nous avons

$$\begin{aligned}
\psi_x &= (\mu_x, \sigma_x^2, \epsilon_x, \mathbf{z}), \\
\psi_y &= (\mu_y, \sigma_y^2, \epsilon_y, \mathbf{w}), \\
f_X(x_{m+1} | \psi_x) &= f_{X,1}(x_{m+1} | \mu_{x_1}, \sigma_{x_1}^2)^{1-z_{m+1}} f_{X,2}(x_{m+1} | \mu_{x_2}, \sigma_{x_2}^2)^{z_{m+1}}, \\
f_Y(y_{n+1} | \psi_y) &= f_{Y,1}(y_{n+1} | \mu_{y_1}, \sigma_{y_1}^2)^{1-w_{n+1}} f_{Y,2}(y_{n+1} | \mu_{y_2}, \sigma_{y_2}^2)^{w_{n+1}},
\end{aligned}$$

où $z_{m+1} \sim \text{Bin}(1, \epsilon_x)$, $w_{n+1} \sim \text{Bin}(1, \epsilon_y)$ et $f_1(\cdot)$ et $f_2(\cdot)$ sont les fonctions de densité d'une loi log-normale.

Pour calculer cette probabilité par simulation Monte Carlo, il s'agit de générer les paramètres ψ_x et ψ_y selon la densité *a posteriori* $\pi_{\psi_x, \psi_y}(\psi_x, \psi_y | \mathbf{x}, \mathbf{y})$, puis à partir de ces paramètres, générer X_{m+1} et Y_{n+1} à partir de la densité $f_{X,Y}(x_{m+1}, y_{n+1} | \psi_x, \psi_y)$. L'algorithme pour calculer numériquement $\mathbb{P}[X_{m+1} < Y_{n+1} | \mathbf{x}, \mathbf{y}]$ sera très semblable à celui du test unilatéral de la section 3.6.

1) Refaire les étapes 1 et 2, soit choisir les paramètres des lois *a priori* et trouver

les estimateurs EM de μ , σ^2 et ϵ .

2) Refaire les étapes 3 à 6 du test unilatéral de la section 3.6, soit générer ϵ_x , ϵ_y , z_i pour $i = 1, \dots, m$, w_i pour $i = 1, \dots, n$, $\sigma_{x_1}^2$, $\sigma_{x_2}^2$, $\sigma_{y_1}^2$, $\sigma_{y_2}^2$, μ_{x_1} , μ_{x_2} , μ_{y_1} et μ_{y_2} .

3) Refaire les deux dernières parties de l'étape 7, soit calculer les fonctions indicatrices $I_X(\mu_{\mathbf{x}}, \sigma_{\mathbf{x}}^2)$ et $I_Y(\mu_{\mathbf{y}}, \sigma_{\mathbf{y}}^2)$.

4) Refaire les étapes 8 à 10, soit calculer $w_{\mathbf{X},1}(\mu_{x_1}, \sigma_{x_1}^2, \epsilon_x, \mathbf{z})$, $w_{\mathbf{X},2}(\mu_{x_2}, \sigma_{x_2}^2, \epsilon_x, \mathbf{z})$, $w_{\mathbf{Y},1}(\mu_{y_1}, \sigma_{y_1}^2, \epsilon_y, \mathbf{w})$, $w_{\mathbf{Y},2}(\mu_{y_2}, \sigma_{y_2}^2, \epsilon_y, \mathbf{w})$ et les constantes $K_{\mathbf{X},3}(\epsilon_x, \mathbf{x})$ et $K_{\mathbf{Y},3}(\epsilon_y, \mathbf{y})$.

5) Calculer

$$\begin{aligned} w_j &= w_{\mathbf{X},1}(\mu_{x_1}, \sigma_{x_1}^2, \epsilon_x, \mathbf{z}) w_{\mathbf{X},2}(\mu_{x_2}, \sigma_{x_2}^2, \epsilon_x, \mathbf{z}) w_{\mathbf{Y},1}(\mu_{y_1}, \sigma_{y_1}^2, \epsilon_y, \mathbf{w}) \\ &\times w_{\mathbf{Y},2}(\mu_{y_2}, \sigma_{y_2}^2, \epsilon_y, \mathbf{w}) K_{\mathbf{X},3}(\epsilon_x, \mathbf{x}) K_{\mathbf{Y},3}(\epsilon_y, \mathbf{y}) I_X(\mu_{\mathbf{x}}, \sigma_{\mathbf{x}}^2) I_Y(\mu_{\mathbf{y}}, \sigma_{\mathbf{y}}^2). \end{aligned}$$

À partir de la prochaine étape, l'algorithme se démarque de celui du test unilatéral de la section 3.6. Il s'agit de générer x_{m+1} et y_{n+1} à partir d'un mélange de lois log-normales avec les paramètres générés dans les étapes précédentes.

6) Générer z_{m+1} selon une loi $Bin(1, \epsilon_x)$, et générer w_{n+1} selon une loi $Bin(1, \epsilon_y)$, où ϵ_x et ϵ_y ont été générées à l'étape 2.

7) Si $z_{m+1} = 1$, générer x_{m+1} selon une loi $LN(\mu_{x_2}, \sigma_{x_2}^2)$,
si $z_{m+1} = 0$, générer x_{m+1} selon une loi $LN(\mu_{x_1}, \sigma_{x_1}^2)$.

8) Si $w_{n+1} = 1$, générer y_{n+1} selon une loi $LN(\mu_{y_2}, \sigma_{y_2}^2)$,
si $w_{n+1} = 0$, générer y_{n+1} selon une loi $LN(\mu_{y_1}, \sigma_{y_1}^2)$.

Notons que nous générons selon une loi $LN(\mu, \sigma^2)$ en générant selon une loi $N(\mu, \sigma^2)$ puis en prenant l'exponentielle.

9) Calculer

$$q_j = \mathbb{I}(x_{m+1} < y_{n+1}) = \begin{cases} 1 & ; \text{ si } x_{m+1} < y_{n+1}, \\ 0 & ; \text{ sinon.} \end{cases}$$

10) Les étapes 2 à 9 représentent une simulation notée j . Refaire ces étapes un grand nombre de fois, disons S , afin de générer une suite w_1, \dots, w_S et une suite q_1, \dots, q_S .

11) Calculer

$$\mathbb{P}[X_{m+1} < Y_{n+1} \mid \mathbf{x}, \mathbf{y}] = \frac{\sum_{j=1}^S q_j w_j}{\sum_{j=1}^S w_j}.$$

Notons que nous pouvons aussi considérer $\sigma_{x_1}^2, \sigma_{x_2}^2, \sigma_{y_1}^2, \sigma_{y_2}^2, \epsilon_x$ et ϵ_y connues. L'algorithme pour les étapes 1 à 5 sera calqué sur celui du test unilatéral de la section 3.4 plutôt que sur celui de la section 3.6, mais restera inchangé pour les étapes 6 à 11.

Nous avons appliqué cet algorithme aux données de l'étude en considérant les paramètres de dispersion et de poids connus et nous trouvons

$$\mathbb{P}[X_{m+1} < Y_{n+1} \mid \mathbf{x}, \mathbf{y}] = 0,764$$

en choisissant les mêmes lois *a priori* que celles choisies à la section 3.8. Nous pouvons donc interpréter ce résultat comme suit: Il y a 76,4% de chance que le coût du traitement 1 soit inférieur au coût du traitement 2 pour une future patiente. Bien que le test d'hypothèse bayésien unilatéral suggérait que $\mathbb{E}(X) < \mathbb{E}(Y)$ avec une probabilité de 0,0044 seulement (voir section 3.8), nous concluons dans cette présente section que $X < Y$ avec une probabilité de 0,764. Cela peut sembler contradictoire, mais cela s'explique par la distribution des données où la moyenne de X est supérieure à celle de Y tandis que la médiane échantillonnale de X est inférieure à celle de Y . Il s'agit ici de bien déterminer la question posée,

à savoir si nous nous intéressons aux coûts totaux à long terme des traitements étudiés (où le traitement 2 est avantage), ou plutôt aux coûts à court terme (où le traitement 1 est avantage).

Dans ce chapitre, nous avons présenté l'approche bayésienne pour comparer deux traitements. La première approche est la méthode de rééchantillonnage paramétrique à partir de modèles ajustés par la méthode bayésienne. Cette méthode permet d'introduire de l'information *a priori* dans une méthode classique qu'est la méthode de rééchantillonnage, afin de répondre au test d'hypothèses d'égalité des moyennes.

La deuxième approche consiste à calculer la probabilité *a posteriori* que l'hypothèse H_0 survienne. Après avoir choisi une fonction de coûts adéquate, nous pouvons prendre, à partir de cette probabilité, une décision entre H_0 et H_1 . Nous avons vu la méthode de calcul numérique par simulations de Monte Carlo avec fonctions d'importance pour estimer cette probabilité. Un algorithme détaillé a été présenté pour les tests unilatéral et bilatéral selon que nous considérons la totalité ou une partie des paramètres comme étant des variables aléatoires. Pour des lois *a priori* très semblables pour les deux traitements, ce qui s'avère un choix logique en pharmacoéconomie, nous arrivons à des résultats assez près de ceux des tests classiques appropriés, soient les méthodes de rééchantillonnage.

Finalement, nous avons abordé une approche autre que les tests d'hypothèses, soit la probabilité *a posteriori* que le premier traitement soit moins coûteux que le deuxième. L'algorithme pour calculer cette probabilité est très semblable à celui utilisé pour le test d'hypothèses unilatéral. Le résultat est toutefois bien loin de celui des tests d'hypothèses, puisque c'est une mesure tout à fait différente pour comparer deux traitements.

CONCLUSION

Dans ce mémoire, nous avons expliqué plusieurs approches statistiques nous permettant de comparer les coûts de deux traitements concurrents. Plusieurs méthodes statistiques sont couramment utilisées en pharmacéconomie pour tester l'égalité des coûts moyens de chaque traitement, dont entre autres le test de Student et le test de Mann-Whitney-Wilcoxon. Ces méthodes ne sont toutefois pas appropriées pour certain types de données. Particulièrement en pharmacéconomie, il arrive fréquemment qu'un médicament engendre des effets secondaires produisant des coûts très élevés pour un petit nombre de patients. La distribution qui en résulte est alors asymétrique, parfois même bimodale avec un mode distinct représentant les coûts des effets secondaires. Nous avons donc présenté des méthodes statistiques alternatives pour comparer adéquatement des traitements avec une telle distribution de coûts.

Nous avons exploré une approche fréquentiste par les méthodes de rééchantillonnage et une approche bayésienne. La méthode de rééchantillonnage non paramétrique a l'avantage d'être relativement simple à appliquer puisqu'on ne suppose aucune distribution paramétrique et elle tient compte de l'asymétrie des données. La méthode de rééchantillonnage paramétrique requiert la modélisation des données, mais a l'avantage de supposer une distribution plus réaliste que la distribution empirique, particulièrement quand la taille des échantillons est petite. Un compromis entre les deux méthodes est le rééchantillonnage modifié où la probabilité d'observer une valeur plus grande que la donnée maximale observée est modélisée par une distribution paramétrique.

Le chapitre 3 est consacré à trois approches bayésiennes. La première est l'application de la méthode de rééchantillonnage paramétrique sur des distributions ajustées aux données par la méthode bayésienne. Cette dernière méthode consiste à trouver les estimateurs, à l'aide de l'algorithme EM, qui maximisent la fonction de densité *a posteriori*. Le test d'hypothèses est effectué selon l'approche fréquentiste mais l'ajustement du modèle est effectué selon l'approche bayésienne.

Une deuxième approche est le test d'hypothèses d'égalité des moyennes selon l'approche bayésienne. Nous avons présenté les tests unilatéral et bilatéral, en supposant une partie ou la totalité des paramètres comme variables aléatoires. Nous avons présenté un algorithme détaillé pour chaque test à la fin de chaque section. Nous calculons par cet algorithme la probabilité que l'hypothèse H_0 survienne. Nous avons appliqué ces algorithmes sur les données de l'étude. En choisissant des lois *a priori* assez semblables pour les deux traitements, nous arrivons à des résultats assez près de ceux de la méthode de rééchantillonnage paramétrique.

Une dernière approche bayésienne présentée est la probabilité *a posteriori* qu'un traitement soit moins coûteux que l'autre. Cette dernière approche est une mesure différente de celles présentées précédemment.

ANNEXE 1. DÉMONSTRATIONS

Démonstration de l'équation (3.4.14)

Développons l'expression suivante afin de démontrer que nous aboutissons à la forme d'une densité d'une loi normale:

$$\begin{aligned}
 & \exp \left[-\frac{1}{2\sigma_{x_2}^2} \sum (\log x_i - \mu_{x_2})^2 z_i \right] \exp \left[-\frac{1}{2\tau_{x_2}^2} (\mu_{x_2} - \lambda_{x_2})^2 \right] \\
 &= \exp \left[-\frac{1}{2\sigma_{x_2}^2} \left(\sum (\log x_i)^2 z_i + \sum \mu_{x_2}^2 z_i - 2 \sum (\log x_i) \mu_{x_2} z_i \right) \right. \\
 &\quad \left. - \frac{1}{2\tau_{x_2}^2} \left(\mu_{x_2}^2 + \lambda_{x_2}^2 - 2\mu_{x_2} \lambda_{x_2} \right) \right] \\
 &= \exp \left[-\frac{1}{2} \left(\mu_{x_2}^2 \left(\frac{\sum z_i}{\sigma_{x_2}^2} + \frac{1}{\tau_{x_2}^2} \right) - 2\mu_{x_2} \left(\frac{\sum z_i \log x_i}{\sigma_{x_2}^2} + \frac{\lambda_{x_2}}{\tau_{x_2}^2} \right) \right. \right. \\
 &\quad \left. \left. + \left(\frac{\sum (\log x_i)^2 z_i}{\sigma_{x_2}^2} + \frac{\lambda_{x_2}^2}{\tau_{x_2}^2} \right) \right) \right] \\
 &\propto \exp \left[-\frac{\tau_{x_2}^2 \sum z_i + \sigma_{x_2}^2}{2\tau_{x_2}^2 \sigma_{x_2}^2} \left(\mu_{x_2} - \frac{\bar{x}_{(2)} \tau_{x_2}^2 \sum z_i + \lambda_{x_2} \sigma_{x_2}^2}{\tau_{x_2}^2 \sum z_i + \sigma_{x_2}^2} \right)^2 \right]
 \end{aligned}$$

où

$$\bar{x}_{(2)} = \frac{\sum_{i=1}^n z_i \log x_i}{\sum_{i=1}^n z_i}.$$

Nous avons donc la forme d'une densité normale pour le paramètre μ_{x_2} .

C.Q.F.D.

Démonstration de l'équation (3.4.22)

Démontrons l'égalité suivante:

$$\begin{aligned}
w_{\mathbf{X},2}(\hat{\mu}_{x_2}, \mathbf{z}) &= \frac{g_{\mathbf{X},2}(\mathbf{x} \mid \mu_{x_2}, \mathbf{z}) \pi_{\mu_{x_2}}(\mu_{x_2}) g_{\mathbf{Z},2}(\mathbf{z})}{\pi_{\mu_{x_2}}(\mu_{x_2} \mid \mathbf{x}, \mathbf{z}) h_{\mathbf{Z},2}(\mathbf{z} \mid \hat{\mu}_{x_2}, \mathbf{x})} \\
&= \frac{\sigma_{x_2}}{\sqrt{\tau_{x_2}^2 \sum z_i + \sigma_{x_2}^2}} \exp \left[\frac{1}{2\sigma_{x_2}^2} (\hat{\mu}_{x_2} - \bar{x}_{(2)})^2 \sum_{i=1}^m z_i \right] \\
&\times \exp \left[-\frac{(\sum_i z_i)(\bar{x}_{(2)} - \lambda_{x_2})^2}{2(\tau_{x_2}^2 \sum_i z_i + \sigma_{x_2}^2)} \right].
\end{aligned}$$

Nous avons

$$\begin{aligned}
&g_{\mathbf{X},2}(\mathbf{x} \mid \mu_{x_2}, \mathbf{z}) \pi_{\mu_{x_2}}(\mu_{x_2}) g_{\mathbf{Z},2}(\mathbf{z}) \\
&= \frac{1}{(\sigma_{x_2})^{\sum_i z_i}} \exp \left[-\frac{1}{2\sigma_{x_2}^2} \sum_{i=1}^m (\log x_i - \mu_{x_2})^2 z_i \right] \\
&\times \frac{1}{\sqrt{2\pi\tau_{x_2}^2}} \exp \left[-\frac{1}{2\tau_{x_2}^2} (\mu_{x_2} - \lambda_{x_2})^2 \right] (\epsilon_x)^{\sum_i z_i} \\
&= \frac{(\epsilon_x)^{\sum_i z_i}}{(\sigma_{x_2})^{\sum_i z_i}} \exp \left[-\frac{1}{2\sigma_{x_2}^2} \sum_{i=1}^m (\log x_i - \bar{x}_{(2)})^2 z_i \right] \\
&\times \exp \left[-\frac{1}{2\sigma_{x_2}^2} \left(\sum_{i=1}^m z_i \right) (\mu_{x_2} - \bar{x}_{(2)})^2 \right] \exp \left[-\frac{1}{2\tau_{x_2}^2} (\mu_{x_2} - \lambda_{x_2})^2 \right] \\
&\times \frac{1}{\sqrt{2\pi\tau_{x_2}^2}} \\
&= \left(\frac{\epsilon_x}{\sigma_{x_2}} \right)^{\sum_i z_i} \exp \left[-\frac{1}{2\sigma_{x_2}^2} \sum_{i=1}^m (\log x_i - \bar{x}_{(2)})^2 z_i \right] \\
&\times \exp \left[-\frac{\tau_{x_2}^2 \sum z_i + \sigma_{x_2}^2}{2\tau_{x_2}^2 \sigma_{x_2}^2} \left(\mu_{x_2} - \left[\frac{\bar{x}_{(2)} \tau_{x_2}^2 \sum z_i + \lambda_{x_2} \sigma_{x_2}^2}{\tau_{x_2}^2 \sum z_i + \sigma_{x_2}^2} \right] \right)^2 \right] \\
&\times \exp \left[-\frac{(\sum_i z_i)(\bar{x}_{(2)} - \lambda_{x_2})^2}{2(\tau_{x_2}^2 \sum_i z_i + \sigma_{x_2}^2)} \right] \frac{1}{\sqrt{2\pi\tau_{x_2}^2}}
\end{aligned}$$

$$\begin{aligned}
&= \left(\frac{\epsilon_x}{\sigma_{x_2}} \right)^{\sum_i z_i} \exp \left[-\frac{1}{2\sigma_{x_2}^2} \sum_{i=1}^m (\log x_i - \hat{\mu}_{x_2})^2 z_i \right] \\
&\times \exp \left[\frac{1}{2\sigma_{x_2}^2} (\hat{\mu}_{x_2} - \bar{x}_{(2)})^2 \sum_{i=1}^m z_i \right] \\
&\times \exp \left[-\frac{\tau_{x_2}^2 \sum z_i + \sigma_{x_2}^2}{2\tau_{x_2}^2 \sigma_{x_2}^2} \left(\mu_{x_2} - \left[\frac{\bar{x}_{(2)} \tau_{x_2}^2 \sum z_i + \lambda_{x_2} \sigma_{x_2}^2}{\tau_{x_2}^2 \sum z_i + \sigma_{x_2}^2} \right] \right)^2 \right] \\
&\times \frac{1}{\sqrt{2\pi\tau_{x_2}^2}} \exp \left[-\frac{(\sum_i z_i)(\bar{x}_{(2)} - \lambda_{x_2})^2}{2(\tau_{x_2}^2 \sum_i z_i + \sigma_{x_2}^2)} \right].
\end{aligned}$$

Nous avons donc

$$\begin{aligned}
g_{\mathbf{X},2}(\mathbf{x} \mid \mu_{x_2}, \mathbf{z}) \pi_{\mu_{x_2}}(\mu_{x_2}) g_{\mathbf{Z},2}(\mathbf{z}) &= h_{\mathbf{Z},2}(\mathbf{z} \mid \hat{\mu}_{x_2}, \mathbf{x}) \\
&\times \exp \left[\frac{1}{2\sigma_{x_2}^2} (\hat{\mu}_{x_2} - \bar{x}_{(2)})^2 \sum_{i=1}^m z_i \right] \\
&\times \pi_{\mu_{x_2}}(\mu_{x_2} \mid \mathbf{x}, \mathbf{z}) \frac{\sigma_{x_2}}{\sqrt{\tau_{x_2}^2 \sum z_i + \sigma_{x_2}^2}} \\
&\times \exp \left[-\frac{(\sum_i z_i)(\bar{x}_{(2)} - \lambda_{x_2})^2}{2(\tau_{x_2}^2 \sum_i z_i + \sigma_{x_2}^2)} \right],
\end{aligned}$$

et finalement,

$$\begin{aligned}
w_{\mathbf{X},2}(\hat{\mu}_{x_2}, \mathbf{z}) &= \frac{g_{\mathbf{X},2}(\mathbf{x} \mid \mu_{x_2}, \mathbf{z}) \pi_{\mu_{x_2}}(\mu_{x_2}) g_{\mathbf{Z},2}(\mathbf{z})}{\pi_{\mu_{x_2}}(\mu_{x_2} \mid \mathbf{x}, \mathbf{z}) h_{\mathbf{Z},2}(\mathbf{z} \mid \hat{\mu}_{x_2}, \mathbf{x})} \\
&= \frac{\sigma_{x_2}}{\sqrt{\tau_{x_2}^2 \sum z_i + \sigma_{x_2}^2}} \exp \left[\frac{1}{2\sigma_{x_2}^2} (\hat{\mu}_{x_2} - \bar{x}_{(2)})^2 \sum_{i=1}^m z_i \right] \\
&\times \exp \left[-\frac{(\sum_i z_i)(\bar{x}_{(2)} - \lambda_{x_2})^2}{2(\tau_{x_2}^2 \sum_i z_i + \sigma_{x_2}^2)} \right].
\end{aligned}$$

C.Q.F.D.

Démonstration de l'équation (3.5.2)

Démontrons que

$$\begin{aligned}
w_{J_2(\mathbf{x}, \mathbf{y})}(\mu_{\mathbf{x}}, \mu_{y_1}, \mathbf{z}, \mathbf{w}) &= w_{\mathbf{X},1}(\hat{\mu}_{x_1}, \mathbf{z}) w_{\mathbf{X},2}(\hat{\mu}_{x_2}, \mathbf{z}) w_{\mathbf{Y},1}(\hat{\mu}_{y_1}, \mathbf{w}) w_{\mathbf{Y},2}(\hat{\mu}_{y_2}, \mu_{y_2}^*, \mathbf{w}) \\
&\times K_{\mathbf{X},1}(\mathbf{x}) K_{\mathbf{X},3}(\mathbf{x}) K_{\mathbf{Y},1}(\mathbf{y}) K_{\mathbf{Y},3}(\mathbf{y}) K_{\mathbf{X},\mathbf{Y}}(\mathbf{x}, \mathbf{y}) \\
&\times I_X(\mu_{x_1}, \mu_{x_2}) I_Y^*(\mu_{y_1}, \mu_{y_2}^*)
\end{aligned}$$

où

$$w'_{Y,2}(\hat{\mu}_{y_2}, \mu_{y_2}^*, \mathbf{w}) = \exp\left(-\frac{1}{2\sigma_{y_2}^2} \sum w_i [(\log y_i - \mu_{y_2}^*)^2 - (\log y_i - \hat{\mu}_{y_2})^2]\right).$$

Comme

$$J_2(\mathbf{x}, \mathbf{y}) = \sum_{\mathbf{z}} \sum_{\mathbf{w}} \int_{\mu_{x_1}} \int_{\mu_{x_2}} \int_{\mu_{y_1}} g(\mu_{x_1}, \mu_{x_2}, \mu_{y_1}, \mu_{y_2}^*, \mathbf{x}, \mathbf{y}, \mathbf{z}, \mathbf{w}) d\mu_{x_1} d\mu_{x_2} d\mu_{y_1}$$

où

$$\begin{aligned} & g(\mu_{x_1}, \mu_{x_2}, \mu_{y_1}, \mu_{y_2}^*, \mathbf{x}, \mathbf{y}, \mathbf{z}, \mathbf{w}) \\ &= K_{X,Y}(\mathbf{x}, \mathbf{y}) f_{X,Y}(\mathbf{x}, \mathbf{y} \mid \mu_{x_1}, \mu_{x_2}, \mu_{y_1}, \mu_{y_2}^*, \mathbf{z}, \mathbf{w}) \pi_{\mu_{x_1}}(\mu_{x_1}) \pi_{\mu_{x_2}}(\mu_{x_2}) \pi_{\mu_{y_1}}(\mu_{y_1}) \\ &\times \pi_{\mathbf{Z}}(\mathbf{z}) \pi_{\mathbf{W}}(\mathbf{w}) I_X(\mu_{x_1}, \mu_{x_2}) I_Y^*(\mu_{y_1}, \mu_{y_2}^*) \\ &= K_{X,Y}(\mathbf{x}, \mathbf{y}) K_{X,1}(\mathbf{x}) g_{X,1}(\mathbf{x} \mid \mu_{x_1}, \mathbf{z}) g_{X,2}(\mathbf{x} \mid \mu_{x_2}, \mathbf{z}) K_{Y,1}(\mathbf{y}) g_{Y,1}(\mathbf{y} \mid \mu_{y_1}, \mathbf{w}) \\ &\times g_{Y,2}(\mathbf{y} \mid \mu_{y_2}^*, \mathbf{w}) \pi_{\mu_{x_1}}(\mu_{x_1}) \pi_{\mu_{x_2}}(\mu_{x_2}) \pi_{\mu_{y_1}}(\mu_{y_1}) g_{Z,1}(\mathbf{z}) g_{Z,2}(\mathbf{z}) g_{W,1}(\mathbf{w}) g_{W,2}(\mathbf{w}) \\ &\times I_X(\mu_{x_1}, \mu_{x_2}) I_Y^*(\mu_{y_1}, \mu_{y_2}^*), \end{aligned}$$

et les fonctions d'importance étant définies comme suit:

$$\begin{aligned} I_{\mu_{\mathbf{x}}, \mu_{y_1}}(\mu_{x_1}, \mu_{x_2}, \mu_{y_1}) &= \pi_{\mu_{x_1}}(\mu_{x_1} \mid \mathbf{x}, \mathbf{z}) \pi_{\mu_{x_2}}(\mu_{x_2} \mid \mathbf{x}, \mathbf{z}) \pi_{\mu_{y_1}}(\mu_{y_1} \mid \mathbf{y}, \mathbf{w}), \\ I_Z(\mathbf{z}) &= \pi_{\mathbf{Z}}(\mathbf{z} \mid \hat{\mu}_{\mathbf{x}}, \mathbf{x}) \\ &= K_{X,3}(\mathbf{x})^{-1} h_{Z,1}(\mathbf{z} \mid \hat{\mu}_{x_1}, \mathbf{x}) h_{Z,2}(\mathbf{z} \mid \hat{\mu}_{x_2}, \mathbf{x}), \\ I_{\mathbf{W}}(\mathbf{w}) &= \pi_{\mathbf{W}}(\mathbf{w} \mid \hat{\mu}_{\mathbf{y}}, \mathbf{y}) \\ &= K_{Y,3}(\mathbf{y})^{-1} h_{W,1}(\mathbf{w} \mid \hat{\mu}_{y_1}, \mathbf{y}) h_{W,2}(\mathbf{w} \mid \hat{\mu}_{y_2}, \mathbf{y}), \end{aligned}$$

alors la fonction de poids est

$$\begin{aligned} w_{J_2(\mathbf{x}, \mathbf{y})}(\mu_{\mathbf{x}}, \mu_{y_1}, \mathbf{z}, \mathbf{w}) &= \frac{g(\mu_{x_1}, \mu_{x_2}, \mu_{y_1}, \mu_{y_2}^*, \mathbf{x}, \mathbf{y}, \mathbf{z}, \mathbf{w})}{I_{\mu_{\mathbf{x}}, \mu_{y_1}}(\mu_{x_1}, \mu_{x_2}, \mu_{y_1}) I_Z(\mathbf{z}) I_{\mathbf{W}}(\mathbf{w})} \\ &= w_{X,1}(\hat{\mu}_{x_1}, \mathbf{z}) w_{X,2}(\hat{\mu}_{x_2}, \mathbf{z}) w_{Y,1}(\hat{\mu}_{y_1}, \mathbf{w}) w'_{Y,2}(\hat{\mu}_{y_2}, \mu_{y_2}^*, \mathbf{w}) \\ &\times K_{X,1}(\mathbf{x}) K_{X,3}(\mathbf{x}) K_{Y,1}(\mathbf{y}) K_{Y,3}(\mathbf{y}) K_{X,Y}(\mathbf{x}, \mathbf{y}) \\ &\times I_X(\mu_{x_1}, \mu_{x_2}) I_Y^*(\mu_{y_1}, \mu_{y_2}^*) \end{aligned}$$

où

$$w_{\mathbf{X},1}(\hat{\mu}_{x_1}, \mathbf{z}) = \frac{g_{\mathbf{X},1}(\mathbf{x} \mid \mu_{x_1}, \mathbf{z}) \pi_{\mu_{x_1}}(\mu_{x_1}) g_{\mathbf{Z},1}(\mathbf{z})}{\pi_{\mu_{x_1}}(\mu_{x_1} \mid \mathbf{x}, \mathbf{z}) h_{\mathbf{Z},1}(\mathbf{z} \mid \hat{\mu}_{x_1}, \mathbf{x})},$$

$$w_{\mathbf{X},2}(\hat{\mu}_{x_2}, \mathbf{z}) = \frac{g_{\mathbf{X},2}(\mathbf{x} \mid \mu_{x_2}, \mathbf{z}) \pi_{\mu_{x_2}}(\mu_{x_2}) g_{\mathbf{Z},2}(\mathbf{z})}{\pi_{\mu_{x_2}}(\mu_{x_2} \mid \mathbf{x}, \mathbf{z}) h_{\mathbf{Z},2}(\mathbf{z} \mid \hat{\mu}_{x_2}, \mathbf{x})},$$

$$w_{\mathbf{Y},1}(\hat{\mu}_{y_1}, \mathbf{w}) = \frac{g_{\mathbf{Y},1}(\mathbf{y} \mid \mu_{y_1}, \mathbf{w}) \pi_{\mu_{y_1}}(\mu_{y_1}) g_{\mathbf{W},1}(\mathbf{w})}{\pi_{\mu_{y_1}}(\mu_{y_1} \mid \mathbf{y}, \mathbf{w}) h_{\mathbf{W},1}(\mathbf{w} \mid \hat{\mu}_{y_1}, \mathbf{y})}$$

et

$$\begin{aligned} w'_{\mathbf{Y},2}(\hat{\mu}_{y_2}, \mu_{y_2}^*, \mathbf{w}) &= \frac{g_{\mathbf{Y},2}(\mathbf{y} \mid \mu_{y_2}^*, \mathbf{w}) g_{\mathbf{W},2}(\mathbf{w})}{h_{\mathbf{W},2}(\mathbf{w} \mid \hat{\mu}_{y_2}, \mathbf{y})} \\ &= \frac{\sigma_{y_2}^{-\sum w_i} \exp \left[-\frac{1}{2\sigma_{y_2}^2} \sum (\log y_i - \mu_{y_2}^*)^2 w_i \right] (\epsilon_y)^{\sum w_i}}{\left(\frac{\epsilon_y}{\sigma_{y_2}} \right)^{\sum w_i} \exp \left[-\frac{1}{2\sigma_{y_2}^2} \sum (\log y_i - \hat{\mu}_{y_2})^2 w_i \right]} \\ &= \exp \left(-\frac{1}{2\sigma_{y_2}^2} \sum w_i \left[(\log y_i - \mu_{y_2}^*)^2 - (\log y_i - \hat{\mu}_{y_2})^2 \right] \right). \end{aligned}$$

C.Q.F.D.

Démonstration de l'équation (3.6.12)

Démontrons l'égalité suivante:

$$\begin{aligned} &w_{\mathbf{X},2}(\mu_{x_2}, \sigma_{x_2}^2, \epsilon_x, \mathbf{z}) \\ &= \frac{g_{\mathbf{X},2}(\mathbf{x} \mid \mu_{x_2}, \sigma_{x_2}^2, \mathbf{z}) \pi_{\mu_{x_2}}(\mu_{x_2}) \pi_{\sigma_{x_2}^2}(\sigma_{x_2}^2) g_{\epsilon_x,2}(\epsilon_x) g_{\mathbf{Z},2}(\mathbf{z} \mid \epsilon_x)}{\pi_{\mu_{x_2}}(\mu_{x_2} \mid \sigma_{x_2}^2, \mathbf{x}, \mathbf{z}) \pi_{\sigma_{x_2}^2}(\sigma_{x_2}^2 \mid \hat{\mu}_{x_2}, \mathbf{x}, \mathbf{z}) h_{\epsilon_x,2}(\epsilon_x \mid \hat{\epsilon}_x) h_{\mathbf{Z},2}(\mathbf{z} \mid \hat{\mu}_{x_2}, \hat{\sigma}_{x_2}^2, \epsilon_x, \mathbf{x})} \\ &= \exp \left[\frac{1}{2\sigma_{x_2}^2} (\hat{\mu}_{x_2} - \bar{x}_{(2)})^2 \sum_{i=1}^m z_i \right] \exp \left[-\frac{(\sum_i z_i) (\bar{x}_{(2)} - \lambda_{x_2})^2}{2(\tau_{x_2}^2 \sum_i z_i + \sigma_{x_2}^2)} \right] \\ &\times \frac{\sigma_{x_2}}{\sqrt{\tau_{x_2}^2 \sum z_i + \sigma_{x_2}^2}} \frac{\beta_{x_2}^{\alpha_{x_2}}}{\Gamma(\alpha_{x_2})} \frac{\Gamma(\sum \frac{z_i}{2} + \alpha_{x_2})}{\left[\frac{1}{2} \sum (\log x_i - \hat{\mu}_{x_2})^2 z_i + \beta_{x_2} \right]^{\sum \frac{z_i}{2} + \alpha_{x_2}}} \\ &\times \left(\frac{\epsilon_x^{m \hat{\epsilon}_x}}{(\hat{\sigma}_{x_2})^{\sum z_i}} \exp \left[-\frac{1}{2\hat{\sigma}_{x_2}^2} (\log x_i - \hat{\mu}_{x_2})^2 z_i \right] \right)^{-1}. \end{aligned}$$

Nous avons les quatres résultats suivants:

i)

$$\begin{aligned} & \frac{g_{X,2}(\mathbf{x} \mid \mu_{x_2}, \sigma_{x_2}^2, \mathbf{z}) \pi_{\mu_{x_2}}(\mu_{x_2})}{\pi_{\mu_{x_2}}(\mu_{x_2} \mid \sigma_{x_2}^2, \mathbf{x}, \mathbf{z})} \\ &= \exp \left[\frac{1}{2\sigma_{x_2}^2} (\hat{\mu}_{x_2} - \bar{x}_{(2)})^2 \sum_{i=1}^m z_i \right] \exp \left[-\frac{(\sum_i z_i)(\bar{x}_{(2)} - \lambda_{x_2})^2}{2(\tau_{x_2}^2 \sum_i z_i + \sigma_{x_2}^2)} \right] \\ &\times \frac{\sigma_{x_2}}{\sqrt{\tau_{x_2}^2 \sum z_i + \sigma_{x_2}^2}} \exp \left[-\frac{1}{2\sigma_{x_2}^2} \sum (\log x_i - \hat{\mu}_{x_2})^2 z_i \right] (\sigma_{x_2})^{-\sum z_i}, \end{aligned}$$

ii)

$$\frac{g_{\epsilon_x,2}(\epsilon_x)}{h_{\epsilon_x,2}(\epsilon_x \mid \hat{\epsilon}_x)} = \epsilon_x^{-m\hat{\epsilon}_x},$$

iii)

$$\begin{aligned} \frac{\pi_{\sigma_{x_2}^2}(\sigma_{x_2}^2)}{\pi_{\sigma_{x_2}^2}(\sigma_{x_2}^2 \mid \hat{\mu}_{x_2}, \mathbf{x}, \mathbf{z})} &= \frac{\beta_{x_2}^{\alpha_{x_2}}}{\Gamma(\alpha_{x_2})} \frac{\Gamma(\sum \frac{z_i}{2} + \alpha_{x_2})}{\left[\frac{1}{2} \sum (\log x_i - \hat{\mu}_{x_2})^2 z_i + \beta_{x_2} \right]^{\sum \frac{z_i}{2} + \alpha_{x_2}}} \\ &\times \frac{(\sigma_{x_2})^{\sum z_i}}{\exp \left[-\frac{1}{2\sigma_{x_2}^2} \sum (\log x_i - \hat{\mu}_{x_2})^2 z_i \right]}, \end{aligned}$$

iv)

$$\frac{g_{Z,2}(\mathbf{z} \mid \epsilon_x)}{h_{Z,2}(\mathbf{z} \mid \hat{\mu}_{x_2}, \hat{\sigma}_{x_2}^2, \epsilon_x, \mathbf{x})} = \left(\exp \left[\frac{-1}{2\hat{\sigma}_{x_2}^2} (\log x_i - \hat{\mu}_{x_2})^2 z_i \right] \right)^{-1} (\hat{\sigma}_{x_2})^{\sum z_i}.$$

Une partie de i) s'annule avec iii), pour le reste, nous retrouvons ces termes dans la fonction de poids.

C.Q.F.D.

Démonstration de l'équation (3.7.1)

Démontrons que

$$\begin{aligned} & w_{J_2(\mathbf{x}, \mathbf{y})}(\mu_{\mathbf{x}}, \mu_{y_1}, \sigma_{\mathbf{x}}^2, \sigma_{y_1}^2, \epsilon_x, \epsilon_y, \mathbf{z}, \mathbf{w}) \\ &= w_{X,1}(\mu_{x_1}, \sigma_{x_1}^2, \epsilon_x, \mathbf{z}) w_{X,2}(\mu_{x_2}, \sigma_{x_2}^2, \epsilon_x, \mathbf{z}) w_{Y,1}(\mu_{y_1}, \sigma_{y_1}^2, \epsilon_y, \mathbf{w}) w'_{Y,2}(\mu_{y_2}^*, \sigma_{y_2}^2, \epsilon_y, \mathbf{w}) \\ &\times K_{X,1}(\mathbf{x}) K_{X,3}(\epsilon_x, \mathbf{x}) K_{X,4}(\mathbf{x}) K_{X,5}(\mathbf{x})^{-1} K_{Y,1}(\mathbf{y}) K_{Y,3}(\epsilon_y, \mathbf{y}) K_{Y,4}(\mathbf{y}) K_{Y,5}(\mathbf{y})^{-1} \\ &\times K_{X,Y}(\mathbf{x}, \mathbf{y}) I_X(\mu_{\mathbf{x}}, \sigma_{\mathbf{x}}^2) I_Y^*(\mu_{y_1}, \mu_{y_2}^*, \sigma_{y_1}^2) \end{aligned}$$

où

$$\begin{aligned}
w'_{Y,2}(\mu_{y_2}^*, \sigma_{y_2}^2, \epsilon_y, \mathbf{w}) &= \exp\left(-\frac{1}{2\sigma_{y_2}^2} \sum w_i [(\log y_i - \mu_{y_2}^*)^2 - (\log y_i - \hat{\mu}_{y_2})^2]\right) \\
&\times \frac{\beta_{y_2}^{\alpha_{y_2}}}{\Gamma(\alpha_{y_2})} \cdot \frac{\Gamma(\frac{\sum w_i}{2} + \alpha_{y_2})}{\left[\frac{1}{2} \sum (\log y_i - \hat{\mu}_{y_2})^2 w_i + \beta_{y_2}\right]^{\frac{\sum w_i}{2} + \alpha_{y_2}}} \\
&\times \left(\frac{\epsilon_y^{n\hat{\epsilon}_y}}{\hat{\sigma}_{y_2}^{\sum w_i}} \exp\left[-\frac{1}{2\hat{\sigma}_{y_2}^2} \sum (\log y_i - \hat{\mu}_{y_2})^2 w_i\right]\right)^{-1}.
\end{aligned}$$

Comme

$$\begin{aligned}
J_2(\mathbf{x}, \mathbf{y}) &= \sum_{\mathbf{z}} \sum_{\mathbf{w}} \int_{\epsilon} \int_{\sigma^2} \int_{\mu_{y_1}} \int_{\mu_{x_2}} \int_{\mu_{x_1}} g(\mu_{x_1}, \mu_{x_2}, \mu_{y_1}, \mu_{y_2}^*, \sigma_{\mathbf{x}}^2, \sigma_{\mathbf{y}}^2, \epsilon_x, \epsilon_y, \mathbf{x}, \mathbf{y}, \mathbf{z}, \mathbf{w}) \\
&\times d\mu_{x_1} d\mu_{x_2} d\mu_{y_1} d\sigma_{\mathbf{x}}^2 d\sigma_{\mathbf{y}}^2 d\epsilon_x d\epsilon_y
\end{aligned}$$

où

$$\begin{aligned}
&g(\mu_{x_1}, \mu_{x_2}, \mu_{y_1}, \mu_{y_2}^*, \sigma_{\mathbf{x}}^2, \sigma_{\mathbf{y}}^2, \epsilon_x, \epsilon_y, \mathbf{x}, \mathbf{y}, \mathbf{z}, \mathbf{w}) \\
&= f_{X,Y}(\mathbf{x}, \mathbf{y} \mid \mu_{\mathbf{x}}, \mu_{y_1}, \mu_{y_2}^*, \sigma^2, \epsilon, \mathbf{z}, \mathbf{w}) \pi_2(\mu_{\mathbf{x}}, \mu_{\mathbf{y}}, \sigma_{\mathbf{x}}^2, \sigma_{\mathbf{y}}^2, \epsilon_x, \epsilon_y) \pi_{\mathbf{Z}}(\mathbf{z}) \pi_{\mathbf{W}}(\mathbf{w}) \\
&= K_{X,Y}(\mathbf{x}, \mathbf{y}) K_{X,1}(\mathbf{x}) g_{X,1}(\mathbf{x} \mid \mu_{x_1}, \sigma_{x_1}^2, \epsilon_x, \mathbf{z}) g_{X,2}(\mathbf{x} \mid \mu_{x_2}, \sigma_{x_2}^2, \epsilon_x, \mathbf{z}) K_{Y,1}(\mathbf{y}) \\
&\times g_{Y,1}(\mathbf{y} \mid \mu_{y_1}, \sigma_{y_1}^2, \epsilon_y, \mathbf{w}) g_{Y,2}(\mathbf{y} \mid \mu_{y_2}^*, \sigma_{y_2}^2, \epsilon_y, \mathbf{w}) \pi_{\mu_{x_1}}(\mu_{x_1}) \pi_{\mu_{x_2}}(\mu_{x_2}) \pi_{\mu_{y_1}}(\mu_{y_1}) \\
&\times \pi_{\sigma_{x_1}^2}(\sigma_{x_1}^2) \pi_{\sigma_{x_2}^2}(\sigma_{x_2}^2) \pi_{\sigma_{y_1}^2}(\sigma_{y_1}^2) \pi_{\sigma_{y_2}^2}(\sigma_{y_2}^2) K_{X,4}(\mathbf{x}) g_{\epsilon_x,1}(\epsilon_x) g_{\epsilon_x,2}(\epsilon_x) K_{Y,4}(\mathbf{y}) g_{\epsilon_y,1}(\epsilon_y) \\
&\times g_{\epsilon_y,2}(\epsilon_y) g_{\mathbf{Z},1}(\mathbf{z}) g_{\mathbf{Z},2}(\mathbf{z}) g_{\mathbf{W},1}(\mathbf{w}) g_{\mathbf{W},2}(\mathbf{w}) I_X(\mu_{\mathbf{x}}, \sigma_{\mathbf{x}}^2) I_Y^*(\mu_{y_1}, \mu_{y_2}^*, \sigma_{\mathbf{y}}^2)
\end{aligned}$$

et les fonctions d'importance étant définies comme suit:

$$\begin{aligned}
I_{\mu_{\mathbf{x}}, \mu_{y_1}}(\mu_{x_1}, \mu_{x_2}, \mu_{y_1}) &= \pi_{\mu_{x_1}}(\mu_{x_1} \mid \sigma_{x_1}^2, \mathbf{x}, \mathbf{z}) \pi_{\mu_{x_2}}(\mu_{x_2} \mid \sigma_{x_2}^2, \mathbf{x}, \mathbf{z}) \pi_{\mu_{y_1}}(\mu_{y_1} \mid \sigma_{y_1}^2, \mathbf{y}, \mathbf{w}), \\
I_{\sigma^2}(\sigma_{x_1}^2, \sigma_{x_2}^2, \sigma_{y_1}^2, \sigma_{y_2}^2) &= \pi_{\sigma_{x_1}^2}(\sigma_{x_1}^2 \mid \hat{\mu}_{x_1}, \mathbf{x}, \mathbf{z}) \pi_{\sigma_{x_2}^2}(\sigma_{x_2}^2 \mid \hat{\mu}_{x_2}, \mathbf{x}, \mathbf{z}) \\
&\times \pi_{\sigma_{y_1}^2}(\sigma_{y_1}^2 \mid \hat{\mu}_{y_1}, \mathbf{y}, \mathbf{w}) \pi_{\sigma_{y_2}^2}(\sigma_{y_2}^2 \mid \hat{\mu}_{y_2}, \mathbf{y}, \mathbf{w}), \\
I_{\epsilon_x}(\epsilon_x) &= K_{X,5}(\mathbf{x}) h_{\epsilon_x,1}(\epsilon_x \mid \hat{\epsilon}_x) h_{\epsilon_x,2}(\epsilon_x \mid \hat{\epsilon}_x), \\
I_{\epsilon_y}(\epsilon_y) &= K_{Y,5}(\mathbf{y}) h_{\epsilon_y,1}(\epsilon_y \mid \hat{\epsilon}_y) h_{\epsilon_y,2}(\epsilon_y \mid \hat{\epsilon}_y), \\
I_{\mathbf{Z}}(\mathbf{z}) &= K_{X,3}(\epsilon_x, \mathbf{x})^{-1} h_{\mathbf{Z},1}(\mathbf{z} \mid \hat{\mu}_{x_1}, \hat{\sigma}_{x_1}^2, \epsilon_x, \mathbf{x}) h_{\mathbf{Z},2}(\mathbf{z} \mid \hat{\mu}_{x_2}, \hat{\sigma}_{x_2}^2, \epsilon_x, \mathbf{x}), \\
I_{\mathbf{W}}(\mathbf{w}) &= K_{Y,3}(\epsilon_y, \mathbf{y})^{-1} h_{\mathbf{W},1}(\mathbf{w} \mid \hat{\mu}_{y_1}, \hat{\sigma}_{y_1}^2, \epsilon_y, \mathbf{y}) h_{\mathbf{W},2}(\mathbf{w} \mid \hat{\mu}_{y_2}, \hat{\sigma}_{y_2}^2, \epsilon_y, \mathbf{y}),
\end{aligned}$$

alors la fonction de poids est

$$\begin{aligned}
& w_{J_2(\mathbf{x}, \mathbf{y})}(\mu_{\mathbf{x}}, \mu_{y_1}, \mathbf{z}, \mathbf{w}) \\
&= \frac{g(\mu_{x_1}, \mu_{x_2}, \mu_{y_1}, \mu_{y_2}^*, \sigma_{x_1}^2, \sigma_{x_2}^2, \sigma_{y_1}^2, \sigma_{y_2}^2, \epsilon_x, \epsilon_y, \mathbf{x}, \mathbf{y}, \mathbf{z}, \mathbf{w})}{I_{\mu_{\mathbf{x}}, \mu_{y_1}}(\mu_{x_1}, \mu_{x_2}, \mu_{y_1}) I_{\sigma^2}(\sigma_{x_1}^2, \sigma_{x_2}^2, \sigma_{y_1}^2, \sigma_{y_2}^2) I_{\epsilon_x}(\epsilon_x) I_{\epsilon_y}(\epsilon_y) I_Z(\mathbf{z}) I_W(\mathbf{w})} \\
&= w_{X,1}(\mu_{x_1}, \sigma_{x_1}^2, \epsilon_x, \mathbf{z}) w_{X,2}(\mu_{x_2}, \sigma_{x_2}^2, \epsilon_x, \mathbf{z}) w_{Y,1}(\mu_{y_1}, \sigma_{y_1}^2, \epsilon_y, \mathbf{w}) \\
&\times w'_{Y,2}(\mu_{y_2}^*, \sigma_{y_2}^2, \epsilon_y, \mathbf{w}) K_{X,1}(\mathbf{x}) K_{X,3}(\mathbf{x}) K_{Y,1}(\mathbf{y}) K_{Y,3}(\mathbf{y}) K_{X,Y}(\mathbf{x}, \mathbf{y}) \\
&\times K_{X,4}(\mathbf{x}) K_{Y,4}(\mathbf{y}) K_{X,5}(\mathbf{x})^{-1} K_{Y,5}(\mathbf{y})^{-1} I_X(\mu_{\mathbf{x}}, \sigma_{\mathbf{x}}^2) I_Y^*(\mu_{y_1}, \mu_{y_2}^*, \sigma_{y_2}^2)
\end{aligned}$$

où

$$\begin{aligned}
w_{X,2}(\mu_{x_2}, \sigma_{x_2}^2, \epsilon_x, \mathbf{z}) &= \\
& \frac{g_{X,2}(\mathbf{x} \mid \mu_{x_2}, \sigma_{x_2}^2, \epsilon_x, \mathbf{z}) \pi_{\mu_{x_2}}(\mu_{x_2}) \pi_{\sigma_{x_2}^2}(\sigma_{x_2}^2) g_{\epsilon_x,2}(\epsilon_x) g_{Z,2}(\mathbf{z})}{\pi_{\mu_{x_2}}(\mu_{x_2} \mid \sigma_{x_2}^2, \mathbf{x}, \mathbf{z}) \pi_{\sigma_{x_2}^2}(\sigma_{x_2}^2 \mid \hat{\mu}_{x_2}, \mathbf{x}, \mathbf{z}) h_{\epsilon_x,2}(\epsilon_x \mid \hat{\epsilon}_x) h_{Z,2}(\mathbf{z} \mid \hat{\mu}_{x_2}, \hat{\sigma}_{x_2}^2, \epsilon_x, \mathbf{x})}
\end{aligned}$$

tel que défini dans le test unilatéral et

$$\begin{aligned}
w'_{Y,2}(\mu_{y_2}^*, \sigma_{y_2}^2, \epsilon_y, \mathbf{w}) &= \frac{g_{Y,2}(\mathbf{y} \mid \mu_{y_2}^*, \sigma_{y_2}^2, \epsilon_y, \mathbf{w}) \pi_{\sigma_{y_2}^2}(\sigma_{y_2}^2) g_{\epsilon_y,2}(\epsilon_y) g_{W,2}(\mathbf{w})}{\pi_{\sigma_{y_2}^2}(\sigma_{y_2}^2 \mid \hat{\mu}_{y_2}, \mathbf{y}, \mathbf{w}) h_{\epsilon_y,2}(\epsilon_y \mid \hat{\epsilon}_y) h_{W,2}(\mathbf{w} \mid \hat{\mu}_{y_2}, \hat{\sigma}_{y_2}^2, \epsilon_y, \mathbf{y})} \\
&= \frac{\sigma_{y_2}^{-\sum w_i} \exp \left[-\frac{1}{2\sigma_{y_2}^2} \sum (\log y_i - \mu_{y_2}^*)^2 w_i \right] \epsilon_y^{\sum w_i}}{\left(\frac{\epsilon_y}{\hat{\sigma}_{y_2}} \right)^{\sum w_i} \exp \left[-\frac{1}{2\hat{\sigma}_{y_2}^2} \sum (\log y_i - \hat{\mu}_{y_2})^2 w_i \right]} \\
&\times \epsilon_y^{-n\hat{\epsilon}_y} \cdot \frac{\beta_{y_2}^{\alpha_{y_2}}}{\Gamma(\alpha_{y_2})} \cdot \frac{\Gamma\left(\frac{\sum w_i}{2} + \alpha_{y_2}\right)}{\left[\frac{1}{2} \sum (\log y_i - \hat{\mu}_{y_2})^2 w_i + \beta_{y_2} \right]^{\frac{\sum w_i}{2} + \alpha_{y_2}}} \\
&\times \frac{(\sigma_{y_2})^{\sum w_i}}{\exp \left[-\frac{1}{2\sigma_{y_2}^2} \sum (\log y_i - \hat{\mu}_{y_2})^2 w_i \right]} \\
&= \exp \left(-\frac{1}{2\sigma_{y_2}^2} \sum w_i \left[(\log y_i - \mu_{y_2}^*)^2 - (\log y_i - \hat{\mu}_{y_2})^2 \right] \right) \\
&\times \frac{\beta_{y_2}^{\alpha_{y_2}}}{\Gamma(\alpha_{y_2})} \cdot \frac{\Gamma\left(\frac{\sum w_i}{2} + \alpha_{y_2}\right)}{\left[\frac{1}{2} \sum (\log y_i - \hat{\mu}_{y_2})^2 w_i + \beta_{y_2} \right]^{\frac{\sum w_i}{2} + \alpha_{y_2}}} \\
&\times \left(\frac{\epsilon_y^{n\hat{\epsilon}_y}}{\hat{\sigma}_{y_2}^{\sum w_i}} \exp \left[-\frac{1}{2\hat{\sigma}_{y_2}^2} \sum (\log y_i - \hat{\mu}_{y_2})^2 w_i \right] \right)^{-1}.
\end{aligned}$$

C.Q.F.D.

ANNEXE 2. PROGRAMMES INFORMATIQUES

```
;*****
; Fichier des fonctions definies pour XLispStat
;*****
;1 Fonction d'arrondissement
;2 Methode de reechantillonnage non parametrique
;3 Methode des moments pour melange de lois (log-)normales
;4 Algorithme EM pour melange de lois (log-)normales
;5 Test de Kolmogorov-Smirnov pour tester une loi
; (log-)normale ou (inverse-)gamma bimodale
;6 Methode de reechantillonnage parametrique
;7 Methode de reechantillonnage modifie
;8 Methode bayesienne d'estimation des parametres
;9 Test unilaterial bayesien par simulations Monte Carlo
;10 Test bilateral bayesien par simulations Monte Carlo
;11 Probabilite a posteriori qu'un traitement soit
;   moins couteux que l'autre

;1 *****
; fonction d'arrondissement
;*****
;entree:
;x=chiffre a arrondir
;pos=nombre de decimales: -1:dizaines 0:unite 1:dixiemes
;*****
;sortie:
;x=chiffre arrondi
;*****
(defun rond (x pos)
  (def mult (^ 10 pos))
  (def x (/ (round (* x mult)) mult))
  x
)
```

```

;2 *****
; Methode de reechantillonnage non parametrique pour
; test d'egalite des moyennes
;*****
;entree:
;x=donnees du 1er echantillon
;y=donnees du 2eme echantillon
;sim=nombre de simulations desire
;*****
;sortie:
;p-value-b: valeur-p du test bilaterale
;p-value-bd: valeur-p du test unilaterale si Ho:ux<uy
;*****
(defun bootstrap (x y sim)
  (def m (length x))
  (def n (length y))
  (def moyvect1 nil)
  (def moyvect2 nil)
  (def moyvect1t nil)
  (def moyvect2t nil)
  (def svect1 nil)
  (def svect2 nil)
  (def svect1t nil)
  (def svect2t nil)

  ;simulation des echantillons et calcul des statistiques
  (def sim1 (truncate (~ sim .5)))
  (def sim2 (round (/ sim sim1)))
  (dotimes (j sim1)
    (dotimes (i sim2)
      (def ech1 (select x (truncate (* m (uniform-rand m)))))
      (def ech2 (select y (truncate (* n (uniform-rand n)))))
      (def moyvect1 (append moyvect1 (list (mean ech1))))
      (def moyvect2 (append moyvect2 (list (mean ech2))))
      (def svect1 (append svect1
        (list (standard-deviation ech1))))
      (def svect2 (append svect2
        (list (standard-deviation ech2))))
    )
    (def moyvect1t (append moyvect1t moyvect1))
    (def moyvect2t (append moyvect2t moyvect2))
    (def svect1t (append svect1t svect1))
    (def svect2t (append svect2t svect2))
    (def moyvect1 nil)

```

```

(def moyvect2 nil)
(def svect1 nil)
(def svect2 nil)
(print (+ j 1))
)

;statistique observe
(def spt (^ (+ (/ (^ (standard-deviation x) 2) m)
  (/ (^ (standard-deviation y) 2) n)) .5))
(def statobs (/ (- (mean x) (mean y)) spt))

;distribution de la statistique
(def spb (^ (+ (/ (^ svect1t 2) m) (/ (^ svect2t 2) n)) .5))
(def diff (/ (- moyvect1t moyvect2t
  (- (mean x) (mean y))) spb))

;calcul de la valeur-p
(def tmpd (select diff (which (> diff (abs statobs))));
(def tmpg (select diff
  (which (< diff (* -1 (abs statobs))));

(def p-value-bd (/ (length tmpd) (* sim1 sim2)))
(def p-value-bg (/ (length tmpg) (* sim1 sim2)))
(def p-value-b (+ p-value-bg p-value-bd))
(print "P-VALUE-BG") (print (rond p-value-bg 4))
(print "P-VALUE-BD") (print (rond p-value-bd 4))
(print "P-VALUE-B") (rond p-value-b 4)
)

;3 *****
; Methode des moments pour melange de lois
; normales ou log-normales bimodales
;*****
;entree:
;x: donnees
;e: epsilon (poids) initial
;u1,u2,s1,s2: parametres initiaux des deux (log-)normales
;logn:=1 si log-normale, 0 si normale
;*****
;sortie:
;e,u1,s1,u2,s2:estimes des parametres du modele
;*****
(defun moment (x e u1 s1 u2 s2 logn iter)
  (if (= logn 1) (def x (log x))

```

```

;estimes initiaux
(def me (- 1 e))
(def v1 (^ s1 2))
(def v2 (^ s2 2))

;moments echantillonnaux
(def x1 (mean x))
(def x2 (mean (^ x 2)))
(def x3 (mean (^ x 3)))
(def x4 (mean (^ x 4)))
(def x5 (mean (^ x 5)))

(dotimes (k iter)
  (def phi (matrix (list 5 1) (list e u1 u2 v1 v2)))
  ;moments theoriques
  (def f1 (- (+ (* me u1) (* e u2)) x1))
  (def f2 (- (+ (* me (+ (^ u1 2) v1))
    (* e (+ (^ u2 2) v2))) x2))
  (def f3 (- (+ (* me (+ (^ u1 3) (* 3 v1 u1)))
    (* e (+ (^ u2 3) (* 3 u2 v2)))) x3))
  (def f4 (- (+ (* me (+ (^ u1 4) (* 6 v1 (^ u1 2))
    (* 3 (^ v1 2)))) (* e
    (+ (^ u2 4) (* 6 v2 (^ u2 2)) (* 3 (^ v2 2)))))) x4))
  (def f5 (- (+ (* me (+ (^ u1 5) (* 10 (^ u1 3) v1)
    (* 15 u1 (^ v1 2)))) (* e (+ (^ u2 5)
    (* 10 (^ u2 3) v2) (* 15 u2 (^ v2 2)))))) x5))
  (def f (matrix (list 5 1) (list f1 f2 f3 f4 f5)))

;derivees
(def d1_e (- u2 u1))
(def d1_u1 me)
(def d1_u2 e)
(def d1_v1 0)
(def d1_v2 0)
(def d2_e (- (+ (^ u2 2) v2) (+ (^ u1 2) v1)))
(def d2_u1 (* 2 me u1))
(def d2_u2 (* 2 e u2))
(def d2_v1 me)
(def d2_v2 e)
(def d3_e (- (+ (^ u2 3) (* 3 u2 v2)) (+ (^ u1 3)
  (* 3 u1 v1))))
(def d3_u2 (* e (+ (* 3 (^ u2 2)) (* 3 v2))))
(def d3_u1 (* me (+ (* 3 (^ u1 2)) (* 3 v1))))
(def d3_v1 (* me 3 u1))

```



```

(def d3_v2 (* e 3 u2))
(def d4_e (- (+ (^ u2 4) (* 6 (^ u2 2) v2) (* 3 (^ v2 2)))
  (+ (^ u1 4) (* 6 (^ u1 2) v1) (* 3 (^ v1 2)))))
(def d4_u1 (* me (+ (* 4 (^ u1 3)) (* 12 v1 u1))))
(def d4_u2 (* e (+ (* 4 (^ u2 3)) (* 12 v2 u2))))
(def d4_v1 (* me (+ (* 6 (^ u1 2)) (* 6 v1))))
(def d4_v2 (* e (+ (* 6 (^ u2 2)) (* 6 v2))))
(def d5_e (- (+ (^ u2 5) (* 10 (^ u2 3) v2)
  (* 15 u2 (^ v2 2)) (+ (^ u1 5)
  (* 10 (^ u1 3) v1) (* 15 u1 (^ v1 2)))))
(def d5_u1 (* me (+ (* 5 (^ u1 4)) (* 30 v1 (^ u1 2))
  (* 15 (^ v1 2)))))
(def d5_u2 (* e (+ (* 5 (^ u2 4)) (* 30 v2 (^ u2 2))
  (* 15 (^ v2 2)))))
(def d5_v1 (* me (+ (* 10 (^ u1 3)) (* 30 u1 v1))))
(def d5_v2 (* e (+ (* 10 (^ u2 3)) (* 30 u2 v2))))

(def g (matrix (list 5 5) (list
  d1_e d1_u1 d1_u2 d1_v1 d1_v2
  d2_e d2_u1 d2_u2 d2_v1 d2_v2
  d3_e d3_u1 d3_u2 d3_v1 d3_v2
  d4_e d4_u1 d4_u2 d4_v1 d4_v2
  d5_e d5_u1 d5_u2 d5_v1 d5_v2
)))
(def ig (inverse g))
(def igf (matmult ig f))
(def phi (- phi igf))
(def e (select phi 0 0))
(def me (- 1 e))
(def u1 (select phi 1 0))
(def u2 (select phi 2 0))
(def v1 (select phi 3 0))
(def v2 (select phi 4 0))
)
(print "e") (print (rond e 3))
(print "u1") (print (rond u1 2))
(print "u2") (print (rond u2 2))
(print "s1") (print (rond (^ v1 .5) 3))
(print "s2") (rond (^ v2 .5) 3)
)

;4 *****
; Algorithmme EM pour melange de lois (log-)normales
;*****

```

```

;entree:
;nb_dec=erreur maximal voulu entre deux estimates
;      du parametre epsilon=10^(-nb_dec)
;y=donnees originales ou log des donnees si log-normales
;p=epsilon initial=poids du deuxieme mode
;*****
;sortie:
;p,u1,s1,u2,s2: estimates des parametres du modele
;stats et p-ks: statistique et valeur-p du test de
;              Kolmogorov-Smirnov
;*****
(defun em_norm (nb_dec y p)
  (def n (length y))
  (def borne (round (* (- 1 p) n)))
  (def u1 (mean (select y (iseq 0 borne))))
  (def u2 (mean (select y (iseq (+ borne 1) (- n 1)))))
  (def s1 (standard-deviation (select y (iseq 0 borne))))
  (def s2 (standard-deviation (select y
    (iseq (+ borne 1) (- n 1)))))
  (def tolerance (^ 10 (- nb_dec)))
  (def cpt 0)
  (def error 1000)
  (loop (if (< error tolerance) (return))
    (def p1 (* (- 1 p) (/ (normal-dens (/ (- y u1) s1)) s1)))
    (def p2 (* p (/ (normal-dens (/ (- y u2) s2)) s2)))
    (def ez (/ p2 (+ p2 p1)))
    (def p_old p)
    (def p (mean ez))
    (def u1 (/ (sum (* y (- 1 ez))) (sum (- 1 ez))))
    (def u2 (/ (sum (* y ez)) (sum ez)))
    (def s1 (^ (/ (sum (* (^ (- y u1) 2) (- 1 ez))
      (sum (- 1 ez))) .5))
    (def s2 (^ (/ (sum (* (^ (- y u2) 2) ez)) (sum ez)) .5))
    (print u1) (print s1)
    (print u2) (print s2)
    (def cpt (+ cpt 1))
    (print p) (print cpt)
    (def error (abs (- p p_old)))
  )
  (def likelihood (vectmult (+ (* (- 1 p)
    (normal-dens (/ (- y u1) s1)))
    (* p (normal-dens (/ (- y u2) s2))))))
  (def u1 (rond u1 6)) (def u2 (rond u2 6))
  (def s1 (rond s1 6)) (def s2 (rond s2 6))

```

```

(def p (rond p 6))
(print "u1") (print (rond u1 2)) (print "s1")
(print (rond s1 3))
(print "u2") (print (rond u2 2)) (print "s2")
(print (rond s2 3))
(print "p") (print (rond p 3))
(print "likelihood") (print likelihood)

;test de Kolmogorov-Smirnov
(def df (+ (* (- 1 p) (normal-cdf (/ (- y u1) s1)))
          (* p      (normal-cdf (/ (- y u2) s2)))))
(def df (sort-data df))
(def statks (* (^ n .5) (max
  (max (abs (- df (/ (iseq 0 (- n 1)) n))))
  (max (abs (- df (/ (iseq 1 n) n)))))))
(def statks (* statks (+ 1 (/ .12 (^ n .5)) (/ .11 n))))
;ajustement car on ne connait pas les parametres de Fo
(print "statks") (print (rond statks 2))
(def p-ks (min 1 (rond (* 2 (exp (* (- 2)
  (^ statks 2)))) 3)))
(print "p-ks") (rond p-ks 3)
)

;5 *****
; Test de Kolmogorov-Smirnov pour melange de lois
; (log-)normales ou (inverse-)gamma
;*****
;entree:
;y=donnees originales
;p1 = alpha ou mu du 1er mode,
;p2 = beta ou ecart type du 1er mode
;p3 = alpha ou mu du 2eme mode,
;p4 = beta ou ecart type du 2eme mode
;epsilon=poids du deuxieme mode
;loi=1 gamma, loi=2 inverse-gamma,
;loi=3 normale, loi=4 log-normale
;*****
;sortie:
;statks et p-ks: statistique et valeur-p du test
;*****

(defun ks_2mode (y p1 p2 p3 p4 epsilon loi)
  (def n (length y))
  (cond

```

```

(= loi 1)
  (def df (+ (* (- 1 epsilon) (gamma-cdf (* y p2) p1))
             (* epsilon (gamma-cdf (* y p4) p3))))
)
(= loi 2)
  (def y (/ 1 y))
  (def df (+ (* (- 1 epsilon) (- 1 (gamma-cdf (* y p2)
                                               p1)))
             (* epsilon (- 1 (gamma-cdf (* y p4) p3)))))
)
(= loi 3)
  (def df (+ (* (- 1 epsilon) (normal-cdf (/ (- y p1) p2)))
             (* epsilon (normal-cdf (/ (- y p3) p4)))))
)
(= loi 4)
  (def y (log y))
  (def df (+ (* (- 1 epsilon) (normal-cdf (/ (- y p1) p2)))
             (* epsilon (normal-cdf (/ (- y p3) p4)))))
)
)
)
(def df (sort-data df))
(def statks (* (^ n .5) (max
  (max (abs (- df (/ (iseq 0 (- n 1)) n))))
  (max (abs (- df (/ (iseq 1 n) n)))))))
)
(def statks (* statks (+ 1 (/ .12 (^ n .5)) (/ .11 n))))
;ajustement car on estime les parametres de la loi
(print "statks") (print statks)
(def p-ks (min 1 (rond (* 2 (exp (* (- 2)
  (^ statks 2)))) 3)))
(print "p-ks") p-ks
)

;6 *****
; Methode de reechantillonnage parametrique pour melange
; de lois (log-)normales
;*****
;entree:
;x: 1er echantillon observe (donnees originales)
;y: 2eme echantillon observe
;sim: nb. de simulations voulues
;ux1: moyenne du 1er mode de la v.a. x
;sx1: ecart type du 1er mode de la v.a. x
;ux2: moyenne du 2eme mode de la v.a. x
;sx2: ecart type du 2eme mode de la v.a. x
;px: pourcentage des donnees associees au 2eme mode

```

```

;      de la v.a. x
;lognx: <>0 si la v.a. x est un melange de log-normales,
;      0 si de normales
;*****
;sortie:
;p-value-s: valeur-p du test bilateral
;p-value-sd: valeur-p du test unilateral si Ho:ux<uy
;*****

(defun bi-normal (x y sim ux1 sx1 ux2 sx2 px lognx
                  uy1 sy1 uy2 sy2 py logny)

  (def m (length x))
  (def n (length y))
  (def moyvect1 nil)
  (def moyvect2 nil)
  (def moyvect1t nil)
  (def moyvect2t nil)
  (def svect1 nil)
  (def svect2 nil)
  (def svect1t nil)
  (def svect2t nil)

;simulation des echantillons
  (def sim1 (truncate (~ sim .5)))
  (def sim2 (round (/ sim sim1)))
  (dotimes (j sim1)
    (dotimes (i sim2)
      (def bx (binomial-rand m 1 px))
      (def ux (+ (* bx ux2) (* (- 1 bx) ux1)))
      (def sx (+ (* bx sx2) (* (- 1 bx) sx1)))
      (if (= lognx 0)
          (def echx (+ (* (normal-rand m) sx) ux))
          (def echx (exp (+ (* (normal-rand m) sx) ux)))
      )
      (def by (binomial-rand n 1 py))
      (def uy (+ (* by uy2) (* (- 1 by) uy1)))
      (def sy (+ (* by sy2) (* (- 1 by) sy1)))
      (if (= logny 0)
          (def echy (+ (* (normal-rand n) sy) uy))
          (def echy (exp (+ (* (normal-rand n) sy) uy)))
      )
      (def moyvect1 (append moyvect1 (list (mean echx))))
      (def moyvect2 (append moyvect2 (list (mean echy))))
      (def svect1 (append svect1

```

```

      (list (standard-deviation echx))))
    (def svect2 (append svect2
      (list (standard-deviation echy))))
  )
  (def moyvect1t (append moyvect1t moyvect1))
  (def moyvect2t (append moyvect2t moyvect2))
  (def svect1t (append svect1t svect1))
  (def svect2t (append svect2t svect2))
  (def moyvect1 nil)
  (def moyvect2 nil)
  (def svect1 nil)
  (def svect2 nil)
  (print (+ j 1))
)

(if (= lognx 0)
  (def mean1 (+ (* (- 1 px) ux1) (* px ux2)))
  (def mean1 (+ (* (- 1 px) (exp (+ ux1 (/ (^ sx1 2) 2))))
    (* px (exp (+ ux2 (/ (^ sx2 2) 2)))))))
)
(if (= logny 0)
  (def mean2 (+ (* (- 1 py) uy1) (* py uy2)))
  (def mean2 (+ (* (- 1 py) (exp (+ uy1 (/ (^ sy1 2) 2))))
    (* py (exp (+ uy2 (/ (^ sy2 2) 2)))))))
)

;statistique observe
(def spt (^ (+ (/ (^ (standard-deviation x) 2) m)
  (/ (^ (standard-deviation y) 2) n)) .5))
(def statobs (/ (- (mean x) (mean y)) spt))

;distribution de la statistique
(def spb (^ (+ (/ (^ svect1t 2) m)
  (/ (^ svect2t 2) n)) .5))
(def diff (/ (- moyvect1t moyvect2t (- mean1 mean2)) spb))

;calcul de la valeur-p
(def tmpd (select diff (which (> diff (abs statobs))));)
(def tmpg (select diff
  (which (< diff (* -1 (abs statobs))));)
(def p-value-sd (/ (length tmpd) (* sim1 sim2)))
(def p-value-sg (/ (length tmpg) (* sim1 sim2)))
(def p-value-s (+ p-value-sg p-value-sd))
(print "P-VALUE-SG")

```

```

(print (rond p-value-sg 4))
(print "P-VALUE-SD")
(print (rond p-value-sd 4))
(print "P-VALUE-S")
(rond p-value-s 4)
)

*;7 *****
; Methode de reechantillonnage modifie avec melange
; de log-normales comme distribution parametrique
;*****
;entree:
;x,y=donnees
;ux2,sx2,px=parametres du 2eme mode de x:mu,sigma et epsilon
;F2x,F2y: =prob d'observer une valeur plus petite
; que la donnee max. pour les x et y selon la distribution
; du 2e mode du melange de lois= F_2(x_{(m)}) et F_2(y_{(n)})
;intx=E(X|X>x_(m))
;sim=nb. de simulations
;*****
;sortie:
;p-value-b: valeur-p du test bilateral
;p-value-bd: valeur-p du test unilateral si Ho:ux<uy
;*****
(defun bootstrap_mod (x y ux2 sx2 px F2x uy2 sy2 py F2y
                    intx inty sim)
  (def m (length x))
  (def n (length y))
  (def lx (* px (- 1 F2x)))
  (def ly (* py (- 1 F2y)))
  (def ex (+ (* (mean x) (- 1 lx)) (* lx intx)))
  (def ey (+ (* (mean y) (- 1 ly)) (* ly inty)))
  (def moyvect1 nil)
  (def moyvect2 nil)
  (def moyvect1t nil)
  (def moyvect2t nil)
  (def svect1 nil)
  (def svect2 nil)
  (def svect1t nil)
  (def svect2t nil)

  ;simulation des echantillons
  (def sim1 (truncate (^ sim .5)))
  (def sim2 (round (/ sim sim1)))

```

```

(dotimes (j sim1)
  (dotimes (i sim2)
    (def unif1 (truncate (* (/ m (- 1 lx)) (uniform-rand m))))
    (def unif1a (select unif1 (which (< unif1 m))))
    (def echx (select x unif1a)) ;###ech. simule
    (cond
      ((> (- m (length echx)) 0)
       (def u (uniform-rand (- m (length echx))))
       (def f (+ (* F2x (- 1 u)) u))
       (def w (round (exp (+ ux2 (* sx2 (normal-quant f))))))
       (def echx (append echx w))
      )
    )
    (def unif2 (truncate (* (/ n (- 1 ly)) (uniform-rand n))))
    (def unif2a (select unif2 (which (< unif2 n))))
    (def echy (select y unif2a)) ;###ech. simule
    (cond
      ((> (- n (length echy)) 0)
       (def u (uniform-rand (- n (length echy))))
       (def f (+ (* F2y (- 1 u)) u))
       (def w (round (exp (+ uy2 (* sy2 (normal-quant f))))))
       (def echy (append echy w))
      )
    )
    (def moyvect1 (append moyvect1 (list (mean echx))))
    (def moyvect2 (append moyvect2 (list (mean echy))))
    (def svect1 (append svect1 (list
      (standard-deviation echx))))
    (def svect2 (append svect2 (list
      (standard-deviation echy))))
    )
    (def moyvect1t (append moyvect1t moyvect1))
    (def moyvect2t (append moyvect2t moyvect2))
    (def svect1t (append svect1t svect1))
    (def svect2t (append svect2t svect2))
    (def moyvect1 nil)
    (def moyvect2 nil)
    (def svect1 nil)
    (def svect2 nil)
    (print (+ j 1))
  )
)

;statistique observe
(def spt (^ (+ (/ (^ (standard-deviation x) 2) m)

```



```

    (/ (^ (standard-deviation y) 2) n)) .5))
(def statobs (/ (- (mean x) (mean y)) spt))

;distribution de la statistique
(def spb (^ (+ (/ (^ svect1t 2) m)
              (/ (^ svect2t 2) n)) .5))
(def diff (/ (- moyvect1t moyvect2t (- ex ey)) spb))

;calcul de la valeur-p
(def tmpd (select diff (which (> diff (abs statobs)))));
(def tmpg (select diff (which
  (< diff (* -1 (abs statobs))))));
(def p-value-bd (/ (length tmpd) (* sim1 sim2)))
(def p-value-bg (/ (length tmpg) (* sim1 sim2)))
(def p-value-b (+ p-value-bg p-value-bd))
(print "P-VALUE-BG") (print (rond p-value-bg 4))
(print "P-VALUE-BD") (print (rond p-value-bd 4))
(print "P-VALUE-B") (rond p-value-b 4)
)

;8 *****
; Algorithmme EM pour melange de lois (log-)normales avec
; lois a priori
;*****
;entree:
;nb_dec=erreur maximal voulu entre deux estimates
;      du parametre epsilon=10^(-nb_dec)
;y=donnees originales ou log des donnees si log-normales
;p=epsilon initial=poids du deuxieme mode
;l1,t1=parametres de la loi a priori sur mu1 = N(l1,t1^2)
;beu,bes=parametres de la loi a priori sur epsilon =
;      Beta t.q. E(Beta)=beu et Ecart type(Be)=bes
;igu1,igs1=parametres de la loi a priori sur sigma^2_1
;      = IG t.q. igu1=E(IG) et igs1=Ecart type(IG)
;*****
;sortie:
;p,u1,s1,u2,s2: estimates des parametres du modele
;statks et p-ks: statistique et valeur-p du test de
;      Kolmogorov-Smirnov
;*****

(defun em_bayes (nb_dec y p beu bes
                l1 t1 l2 t2 igu1 igs1 igu2 igs2)
  (def t1c (^ t1 2))

```

```

(def t2c (^ t2 2))
(def a1 (+ 2 (/ (^ igu1 2) (^ igs1 2))))
(def b1 (* igu1 (- a1 1)))
(def a2 (+ 2 (/ (^ igu2 2) (^ igs2 2))))
(def b2 (* igu2 (- a2 1)))
(def c (/ (- (* (- 1 beu) (^ beu 2))
  (* beu (^ bes 2))) (^ bes 2)))
(def d (* c (/ (- 1 beu) beu)))
(def n (length y))
(def borne (round (* (- 1 p) n)))
(def u1 (mean (select y (iseq 0 borne))))
(def u2 (mean (select y (iseq (+ borne 1) (- n 1)))))
(def s1 (standard-deviation (select y (iseq 0 borne))))
(def s2 (standard-deviation (select y
  (iseq (+ borne 1) (- n 1)))))
(def s1c (^ s1 2))
(def s2c (^ s2 2))
(def tolerance (^ 10 (- nb_dec)))
(def cpt 0)
(def error 1000)
(loop (if (< error tolerance) (return))
  (def p1 (* (- 1 p) (/ (normal-dens (/ (- y u1) s1)) s1)))
  (def p2 (* p (/ (normal-dens (/ (- y u2) s2)) s2)))
  (def ez (/ p2 (+ p2 p1)))
  (def p_old p)
  (def p (/ (+ (sum ez) c (- 1)) (+ n c d (- 2))))
  (def u1 (/ (+ (* t1c (sum (* y (- 1 ez)))) (* s1c 11))
    (+ s1c (* t1c (sum (- 1 ez))))))
  (def u2 (/ (+ (* t2c (sum (* y ez))) (* s2c 12))
    (+ s2c (* t2c (sum ez)))))
  (def s1 (^ (/ (+ b1 (/ (sum (* (^ (- y u1) 2) (- 1 ez))) 2))
    (+ 1 a1 (* .5 (sum (- 1 ez)))) .5))
  (def s2 (^ (/ (+ b2 (/ (sum (* (^ (- y u2) 2) ez)) 2))
    (+ 1 a2 (* .5 (sum ez)))) .5))
  (print u1) (print s1)
  (print u2) (print s2)
  (def cpt (+ cpt 1))
  (print p) (print cpt)
  (def error (abs (- p p_old)))
)
(def likelihood (vectmult (+ (* (- 1 p)
  (normal-dens (/ (- y u1) s1)))
  (* p (normal-dens (/ (- y u2) s2)))))
(def u1 (rond u1 6)) (def u2 (rond u2 6))

```

```

(def s1 (rond s1 6)) (def s2 (rond s2 6))
(def p (rond p 6))
(print "u1") (print (rond u1 2)) (print "s1")
(print (rond s1 3))
(print "u2") (print (rond u2 2)) (print "s2")
(print (rond s2 3))
(print "p") (print (rond p 3))
(print "likelihood") (print likelihood)

;test de Kolmogorov-Smirnov
(def df (+ (* (- 1 p) (normal-cdf (/ (- y u1) s1)))
          (* p      (normal-cdf (/ (- y u2) s2)))))
(def df (sort-data df))
(def statks (* (^ n .5) (max
  (max (abs (- df (/ (iseq 0 (- n 1) n))))
        (max (abs (- df (/ (iseq 1 n) n)))))))
(def statks (* statks (+ 1 (/ .12 (^ n .5)) (/ .11 n))))
;ajustement car on ne connait pas les parametres de Fo
(print "statks") (print (rond statks 2))
(def p-ks (min 1 (rond (* 2 (exp (* (- 2)
  (^ statks 2)))) 3)))
(print "p-ks") (rond p-ks 3)
)

;9 *****
; test d'hypothese bayesien unilateral par Monte Carlo
; avec sigma et epsilon fixes
;*****
;entree:
;x,y=donnees originales
;sim=nb. de simulations
;lx1,lx2=parametres de la loi a priori sur
; mux1 ~ N(lx1,tx1^2) et mux2 ~ N(lx2,tx2^2)
;px1,px2,py1,py2=constantes pour stabiliser les fonctions
; de poids
;*****
;sortie:
;prob=Prob[Ho:mux<muy|x,y]
;*****

(defun unimcarlo1 (x y sim lx1 lx2 ly1 ly2 tx1 tx2 ty1 ty2
  px1 px2 py1 py2)
  (def nx (length x))
  (def ny (length y))

```

```

(def t2x1 (^ tx1 2))
(def t2x2 (^ tx2 2))
(def t2y1 (^ ty1 2))
(def t2y2 (^ ty2 2))
(def ux1c 5.47)
(def ux2c 7.58)
(def uy1c 5.57)
(def uy2c 6.05) ;par ajustement EM
(def sx1 .363)
(def sx2 .558)
(def sy1 .169)
(def sy2 .323) ;par ajustement EM
(def s2x1 (^ sx1 2))
(def s2x2 (^ sx2 2))
(def s2y1 (^ sy1 2))
(def s2y2 (^ sy2 2))
(def px 0.162)
(def py 0.139) ;par ajustement EM
(def ho 0)
(def h1 0)

(def z1 (* (- 1 px) (/ (normal-dens (/ (- (log x) ux1c)
  sx1)) sx1)))
(def z2 (* px (/ (normal-dens (/ (- (log x) ux2c) sx2))
  sx2)))
(def zm (/ z2 (+ z1 z2)))

(def w1 (* (- 1 py) (/ (normal-dens (/ (- (log y) uy1c)
  sy1)) sy1)))
(def w2 (* py (/ (normal-dens (/ (- (log y) uy2c) sy2))
  sy2)))
(def wm (/ w2 (+ w1 w2)))

(dotimes (k sim)
  (def w (select (transpose (binomial-rand 1 1 wm)) 0))
  (def yb1 (/ (sum (* (log y) (- 1 w))) (sum (- 1 w))))
  (def yb2 (/ (sum (* (log y) w)) (sum w)))
  (def muy1 (/ (+ (* yb1 t2y1) (* ly1 s2y1
    (/ (sum (- 1 w)))) (+ t2y1 (/ s2y1 (sum (- 1 w))))))
  (def muy2 (/ (+ (* yb2 t2y2) (* ly2 s2y2 (/ (sum w)))
    (+ t2y2 (/ s2y2 (sum w)))))
  (def vuy1 (/ (* t2y1 s2y1 (/ (sum (- 1 w))))
    (+ t2y1 (/ s2y1 (sum (- 1 w)))))
  (def vuy2 (/ (* t2y2 s2y2 (/ (sum w))))

```

```

(+ t2y2 (/ s2y2 (sum w))))
(def uy1 (select (+ (* (normal-rand 1) (^ vuy1 .5))
  muy1) 0))
(def uy2 (select (+ (* (normal-rand 1) (^ vuy2 .5))
  muy2) 0))
(def poidsy1 (* (^ vuy1 .5) (exp (* (^ (- uy1c yb1) 2)
  (sum (- 1 w)) (/ .5 s2y1)))) (exp (/ (* (sum (- 1 w))
  (^ (- yb1 ly1) 2)) (* -2 (+ s2y1 (* t2y1
  (sum (- 1 w))))))))))
(def poidsy2 (* (^ vuy2 .5) (exp (* (^ (- uy2c yb2) 2)
  (sum w) (/ .5 s2y2)))) (exp (/ (* (sum w)
  (^ (- yb2 ly2) 2)) (* -2 (+ s2y2 (* t2y2 (sum w))))))))))
(def poidsy1 (/ poidsy1 py1))
(def poidsy2 (/ poidsy2 py2))
(def ey1 (exp (+ uy1 (/ s2y1 2))))
(def ey2 (exp (+ uy2 (/ s2y2 2))))
(def ey (+ (* (- 1 py) ey1) (* py ey2)))

(def z (select (transpose (binomial-rand 1 1 zm)) 0))
(def xb1 (/ (sum (* (log x) (- 1 z))) (sum (- 1 z))))
(def xb2 (/ (sum (* (log x) z)) (sum z)))
(def mux1 (/ (+ (* xb1 t2x1) (* lx1 s2x1
  (/ (sum (- 1 z)))))) (+ t2x1 (/ s2x1 (sum (- 1 z))))))
(def mux2 (/ (+ (* xb2 t2x2) (* lx2 s2x2 (/ (sum z))))
  (+ t2x2 (/ s2x2 (sum z))))))
(def vux1 (/ (* t2x1 s2x1 (/ (sum (- 1 z))))
  (+ t2x1 (/ s2x1 (sum (- 1 z))))))
(def vux2 (/ (* t2x2 s2x2 (/ (sum z)))
  (+ t2x2 (/ s2x2 (sum z))))))
(def ux1 (select (+ (* (normal-rand 1) (^ vux1 .5))
  mux1) 0))
(def ux2 (select (+ (* (normal-rand 1) (^ vux2 .5))
  mux2) 0))
(def poidsx1 (* (^ vux1 .5) (exp (* (^ (- ux1c xb1) 2)
  (sum (- 1 z)) (/ .5 s2x1)))) (exp (/ (* (sum (- 1 z))
  (^ (- xb1 lx1) 2)) (* -2 (+ s2x1 (* t2x1
  (sum (- 1 z))))))))))
(def poidsx2 (* (^ vux2 .5) (exp (* (^ (- ux2c xb2) 2)
  (sum z) (/ .5 s2x2)))) (exp (/ (* (sum z)
  (^ (- xb2 lx2) 2)) (* -2 (+ s2x2 (* t2x2 (sum z))))))))))
(def poidsx1 (/ poidsx1 px1))
(def poidsx2 (/ poidsx2 px2))
(def ex1 (exp (+ ux1 (/ s2x1 2))))
(def ex2 (exp (+ ux2 (/ s2x2 2))))

```

```

(def ex (+ (* (- 1 px) ex1) (* px ex2)))

(def poids_d (* poidsx1 poidsx2 poidsy1 poidsy2))

(cond ((> ex1 ex2) NIL)
      ((> ey1 ey2) NIL)
      ((< ex ey) (def ho (+ ho poids_d)
                      (def h1 (+ h1 poids_d)))
      ((>= ex ey) (def h1 (+ h1 poids_d)))
      )
; (print k)
)
(def prob (/ ho h1))
(print "ho") (print ho)
(print "h1") (print h1)
(print "valeur-p") prob
; (print poidsx1)
; (print poidsx2)
; (print poidsy1)
; poidsy2
)

;10 *****
; test d'hypothese bayesien bilaterale par Monte Carlo
; avec sigma et epsilon fixes
; *****
;entree:
;x,y=donnees originales
;sim=nb. de simulations
;p0=Prob[Ho:mux=muy] a priori
;lx1,lx2=parametres de la loi a priori sur
; mux1 ~ N(lx1,tx1^2) et mux2 ~ N(lx2,tx2^2)
;px1,px2,py1,py2=constantes pour stabiliser les fonctions
; de poids
; *****
;sortie:
;prob=Prob[Ho:mux=muy|x,y]
; *****

(defun bimcarlo1 (x y sim p0 lx1 lx2 ly1 ly2 tx1 tx2
                 ty1 ty2 px1 px2 py1 py2)
  (def nx (length x))
  (def ny (length y))
  (def t2x1 (^ tx1 2))

```

```

(def t2x2 (^ tx2 2))
(def t2y1 (^ ty1 2))
(def t2y2 (^ ty2 2))
(def ux1c 5.47)
(def ux2c 7.58)
(def uy1c 5.57)
(def uy2c 6.05) ;par ajustement EM
(def sx1 .363)
(def sx2 .558)
(def sy1 .169)
(def sy2 .323) ;par ajustement EM
(def s2x1 (^ sx1 2))
(def s2x2 (^ sx2 2))
(def s2y1 (^ sy1 2))
(def s2y2 (^ sy2 2))
(def px 0.162)
(def py 0.139) ;par ajustement EM
(def ho 0)
(def h1 0)

(def z1 (* (- 1 px) (/ (normal-dens (/ (- (log x) ux1c)
  sx1)) sx1)))
(def z2 (* px (/ (normal-dens (/ (- (log x) ux2c) sx2))
  sx2)))
(def zm (/ z2 (+ z1 z2)))

(def w1 (* (- 1 py) (/ (normal-dens (/ (- (log y) uy1c)
  sy1)) sy1)))
(def w2 (* py (/ (normal-dens (/ (- (log y) uy2c) sy2))
  sy2)))
(def wm (/ w2 (+ w1 w2)))

(dotimes (k sim)
  (def w (select (transpose (binomial-rand 1 1 wm)) 0))
  (def yb1 (/ (sum (* (log y) (- 1 w)) (sum (- 1 w))))
  (def yb2 (/ (sum (* (log y) w) (sum w)))
  (def muy1 (/ (+ (* yb1 t2y1) (* ly1 s2y1
    (/ (sum (- 1 w)))) (+ t2y1 (/ s2y1 (sum (- 1 w))))))
  (def muy2 (/ (+ (* yb2 t2y2) (* ly2 s2y2 (/ (sum w)))
    (+ t2y2 (/ s2y2 (sum w)))))
  (def vuy1 (/ (* t2y1 s2y1 (/ (sum (- 1 w))))
    (+ t2y1 (/ s2y1 (sum (- 1 w)))))
  (def vuy2 (/ (* t2y2 s2y2 (/ (sum w))
    (+ t2y2 (/ s2y2 (sum w)))))

```

```

(def uy1 (select (+ (* (normal-rand 1) (^ vuy1 .5))
  muy1) 0))
(def uy2 (select (+ (* (normal-rand 1) (^ vuy2 .5))
  muy2) 0))
(def poidsy1 (* (^ vuy1 .5) (exp (* (^ (- uy1c yb1) 2)
  (sum (- 1 w)) (/ .5 s2y1))) (exp (/ (* (sum (- 1 w))
  (^ (- yb1 ly1) 2)) (* -2 (+ s2y1 (* t2y1
  (sum (- 1 w))))))))))
(def poidsy2 (* (^ vuy2 .5) (exp (* (^ (- uy2c yb2) 2)
  (sum w) (/ .5 s2y2))) (exp (/ (* (sum w)
  (^ (- yb2 ly2) 2)) (* -2 (+ s2y2 (* t2y2 (sum w))))))))))
(def poidsy1 (/ poidsy1 py1))
(def poidsy2 (/ poidsy2 py2))
(def ey1 (exp (+ uy1 (/ s2y1 2))))
(def ey2 (exp (+ uy2 (/ s2y2 2))))
(def ey (+ (* (- 1 py) ey1) (* py ey2)))

(def z (select (transpose (binomial-rand 1 1 zm)) 0))
(def xb1 (/ (sum (* (log x) (- 1 z))) (sum (- 1 z))))
(def xb2 (/ (sum (* (log x) z)) (sum z)))
(def mux1 (/ (+ (* xb1 t2x1) (* lx1 s2x1
  (/ (sum (- 1 z)))) (+ t2x1 (/ s2x1 (sum (- 1 z))))))
(def mux2 (/ (+ (* xb2 t2x2) (* lx2 s2x2 (/ (sum z)))
  (+ t2x2 (/ s2x2 (sum z))))))
(def vux1 (/ (* t2x1 s2x1 (/ (sum (- 1 z))))
  (+ t2x1 (/ s2x1 (sum (- 1 z))))))
(def vux2 (/ (* t2x2 s2x2 (/ (sum z)))
  (+ t2x2 (/ s2x2 (sum z))))))
(def ux1 (select (+ (* (normal-rand 1) (^ vux1 .5))
  mux1) 0))
(def ux2 (select (+ (* (normal-rand 1) (^ vux2 .5))
  mux2) 0))
(def poidsx1 (* (^ vux1 .5) (exp (* (^ (- ux1c xb1) 2)
  (sum (- 1 z)) (/ .5 s2x1))) (exp (/ (* (sum (- 1 z))
  (^ (- xb1 lx1) 2)) (* -2 (+ s2x1 (* t2x1
  (sum (- 1 z))))))))))
(def poidsx2 (* (^ vux2 .5) (exp (* (^ (- ux2c xb2) 2)
  (sum z) (/ .5 s2x2))) (exp (/ (* (sum z)
  (^ (- xb2 lx2) 2)) (* -2 (+ s2x2 (* t2x2 (sum z))))))))))
(def poidsx1 (/ poidsx1 px1))
(def poidsx2 (/ poidsx2 px2))
(def ex1 (exp (+ ux1 (/ s2x1 2))))
(def ex2 (exp (+ ux2 (/ s2x2 2))))
(def ex (+ (* (- 1 px) ex1) (* px ex2)))

```



```

(def ey2f (/ (- ex (* (- 1 py) ey1)) py))
(def uy2f (- (log ey2f) (/ s2y2 2)))
(def poidsy2b (exp (* (/ -.5 s2y2) (sum (* w (- (-
  (- (log y) uy2f) 2) (- (- (log y) uy2c) 2)))))))
(def poidsy2b (/ poidsy2b py2))

(def poids_n (* poidsx1 poidsx2 poidsy1 poidsy2b))
(def poids_d (* poidsx1 poidsx2 poidsy1 poidsy2))

(cond ((> ex1 ex2) NIL)
      ((> ey1 ey2) NIL)
      ((> ey1 ey2f) NIL)
      ((def ho (+ ho poids_n))
       (def h1 (+ h1 poids_d)) )
      )
(print k)
)
(def r (/ ho h1))
(def prob (/ (+ 1 (/ (- 1 p0) (* p0 r))))))
(print "ho") (print ho)
(print "h1") (print h1)
(print "r") (print r)
(print "valeur-p") prob
; (print poidsx1)
; (print poidsx2)
; (print poidsy1)
; (print poidsy2)
; poidsy2b
)

;11 *****
; probabilite que le traitement 1 soit moins couteux
; que le traitement 2 avec sigma et epsilon fixes
;*****
;entree:
;x,y=donnees originales
;sim=nb. de simulations
;lx1,lx2=parametres de la loi a priori sur
; mux1 ~ N(lx1,tx1^2) et mux2 ~ N(lx2,tx2^2)
;px1,px2,py1,py2=constantes pour stabiliser les fonctions
; de poids
;*****
;sortie:

```

```

;prob=Prob[cout du traitement 1<cout du traitement 2]
;*****

(defun pxymcarlo1 (x y sim lx1 lx2 ly1 ly2 tx1 tx2 ty1 ty2
                  px1 px2 py1 py2)
  (def nx (length x))
  (def ny (length y))
  (def t2x1 (^ tx1 2))
  (def t2x2 (^ tx2 2))
  (def t2y1 (^ ty1 2))
  (def t2y2 (^ ty2 2))
  (def ux1c 5.47)
  (def ux2c 7.58)
  (def uy1c 5.57)
  (def uy2c 6.05) ;par ajustement EM
  (def sx1 .363)
  (def sx2 .558)
  (def sy1 .169)
  (def sy2 .323) ;par ajustement EM
  (def s2x1 (^ sx1 2))
  (def s2x2 (^ sx2 2))
  (def s2y1 (^ sy1 2))
  (def s2y2 (^ sy2 2))
  (def px 0.162)
  (def py 0.139) ;par ajustement EM
  (def ho 0)
  (def h1 0)

  (def z1 (* (- 1 px) (/ (normal-dens (/ (- (log x) ux1c)
                                         sx1)) sx1)))
  (def z2 (* px (/ (normal-dens (/ (- (log x) ux2c)
                                         sx2))
                    sx2)))
  (def zm (/ z2 (+ z1 z2)))

  (def w1 (* (- 1 py) (/ (normal-dens (/ (- (log y) uy1c)
                                         sy1)) sy1)))
  (def w2 (* py (/ (normal-dens (/ (- (log y) uy2c)
                                         sy2))
                    sy2)))
  (def wm (/ w2 (+ w1 w2)))

  (dotimes (k sim)
    (def w (select (transpose (binomial-rand 1 1 wm)) 0))
    (def yb1 (/ (sum (* (log y) (- 1 w))) (sum (- 1 w))))
    (def yb2 (/ (sum (* (log y) w)) (sum w)))

```

```

(def muy1 (/ (+ (* yb1 t2y1) (* ly1 s2y1
  (/ (sum (- 1 w)))))) (+ t2y1 (/ s2y1 (sum (- 1 w))))))
(def muy2 (/ (+ (* yb2 t2y2) (* ly2 s2y2 (/ (sum w))))
  (+ t2y2 (/ s2y2 (sum w))))))
(def vuy1 (/ (* t2y1 s2y1 (/ (sum (- 1 w))))
  (+ t2y1 (/ s2y1 (sum (- 1 w))))))
(def vuy2 (/ (* t2y2 s2y2 (/ (sum w)))
  (+ t2y2 (/ s2y2 (sum w))))))
(def uy1 (select (+ (* (normal-rand 1) (^ vuy1 .5))
  muy1) 0))
(def uy2 (select (+ (* (normal-rand 1) (^ vuy2 .5))
  muy2) 0))
(def poidsy1 (* (^ vuy1 .5) (exp (* (^ (- uy1c yb1) 2)
  (sum (- 1 w)) (/ .5 s2y1)))) (exp (/ (* (sum (- 1 w))
  (^ (- yb1 ly1) 2)) (* -2 (+ s2y1 (* t2y1
  (sum (- 1 w))))))))))
(def poidsy2 (* (^ vuy2 .5) (exp (* (^ (- uy2c yb2) 2)
  (sum w) (/ .5 s2y2)))) (exp (/ (* (sum w)
  (^ (- yb2 ly2) 2)) (* -2 (+ s2y2 (* t2y2 (sum w))))))))))
(def poidsy1 (/ poidsy1 py1))
(def poidsy2 (/ poidsy2 py2))
(def ey1 (exp (+ uy1 (/ s2y1 2))))
(def ey2 (exp (+ uy2 (/ s2y2 2))))
(def ey (+ (* (- 1 py) ey1) (* py ey2)))

(def z (select (transpose (binomial-rand 1 1 zm)) 0))
(def xb1 (/ (sum (* (log x) (- 1 z))) (sum (- 1 z))))
(def xb2 (/ (sum (* (log x) z)) (sum z)))
(def mux1 (/ (+ (* xb1 t2x1) (* lx1 s2x1
  (/ (sum (- 1 z)))))) (+ t2x1 (/ s2x1 (sum (- 1 z))))))
(def mux2 (/ (+ (* xb2 t2x2) (* lx2 s2x2 (/ (sum z))))
  (+ t2x2 (/ s2x2 (sum z))))))
(def vux1 (/ (* t2x1 s2x1 (/ (sum (- 1 z))))
  (+ t2x1 (/ s2x1 (sum (- 1 z))))))
(def vux2 (/ (* t2x2 s2x2 (/ (sum z)))
  (+ t2x2 (/ s2x2 (sum z))))))
(def ux1 (select (+ (* (normal-rand 1) (^ vux1 .5))
  mux1) 0))
(def ux2 (select (+ (* (normal-rand 1) (^ vux2 .5))
  mux2) 0))
(def poidsx1 (* (^ vux1 .5) (exp (* (^ (- ux1c xb1) 2)
  (sum (- 1 z)) (/ .5 s2x1)))) (exp (/ (* (sum (- 1 z))
  (^ (- xb1 lx1) 2)) (* -2 (+ s2x1 (* t2x1
  (sum (- 1 z))))))))))

```

```

(def poidsx2 (* (^ vux2 .5) (exp (* (^ (- ux2c xb2) 2)
  (sum z) (/ .5 s2x2)))) (exp (/ (* (sum z)
  (^ (- xb2 lx2) 2)) (* -2 (+ s2x2 (* t2x2 (sum z))))))))
(def poidsx1 (/ poidsx1 px1))
(def poidsx2 (/ poidsx2 px2))
(def ex1 (exp (+ ux1 (/ s2x1 2))))
(def ex2 (exp (+ ux2 (/ s2x2 2))))
(def ex (+ (* (- 1 px) ex1) (* px ex2)))

(def zm1 (select (binomial-rand 1 1 px) 0))
(def xm1 (exp (+ (* (- 1 zm1) (+ ux1 (* (select
  (normal-rand 1) 0) s2x1)))) (* zm1 (+ ux2 (* (select
  (normal-rand 1) 0) s2x2))))))
(def wn1 (select (binomial-rand 1 1 py) 0))
(def yn1 (exp (+ (* (- 1 wn1) (+ uy1 (* (select
  (normal-rand 1) 0) s2y1)))) (* wn1 (+ uy2 (* (select
  (normal-rand 1) 0) s2y2))))))

(def poids_d (* poidsx1 poidsx2 poidsy1 poidsy2))

(cond ((> ex1 ex2) NIL)
      ((> ey1 ey2) NIL)
      ((< xm1 yn1) (def ho (+ ho poids_d))
              (def h1 (+ h1 poids_d))
              ((>= xm1 yn1) (def h1 (+ h1 poids_d)))
      )
; (print k)
)
(def prob (/ ho h1))
(print "ho") (print ho)
(print "h1") (print h1)
(print "prob.") prob
; (print poidsx1)
; (print poidsx2)
; (print poidsy1)
; poidsy2
)

```

RÉFÉRENCES

- Bickel, P.J. et Doksum, K.A. (1977), *Mathematical statistics: basic ideas and selected topics*, Holden-Day, Oakland.
- Box, G.E.P., Cox, D.R. (1964), An analysis of transformation, *Journal of the Royal Statistical Society, Series B*, **26**, 211.
- Cohen, A.C. (1967), Estimation in mixtures of two normal distributions, *Technometrics*, **9**, 15-28.
- Coyle, D. (1996), Statistical analysis in pharmacoeconomic studies: a review of current issues and standards, *Pharmacoeconomics*, **9 (6)**, 506-16.
- Daniel, W.W. (1990), *Applied nonparametric statistics*, Pws-Kent, Boston.
- Desgagné, A., Castilloux, A.-M., Angers, J.-F., LeLorier, J. (1998), The use of the bootstrap statistical method for the pharmacoeconomic cost analysis of skewed data, *Pharmacoeconomics*, **13 (5 Pt 1)**, 487-97.
- Dempster, A.P., Laird, N.M., Rubin, D.B. (1977), Maximum likelihood from incomplete data via the EM algorithm, *Journal of the Royal Statistical Society, Series B*, **39**, 1-38.
- Efron, B. et Tibshirani, R.J. (1993), *An introduction to the bootstrap*, Chapman and Hall, New York.
- Hall, P. (1992), *The bootstrap and Edgeworth expansion*, Springer-Verlag, New York.
- Hogg, R.V. et Tanis, E.A. (1988), *Probability and statistical inference*, Macmillan, New York.
- Randles, R.H. et Wolfe, D.A. (1979), *Introduction to the theory of nonparametric*

statistics, John Wiley, New York.

Rao, C.R. (1948), The utilization of multiple measurements in problems of biological classification, *Journal of the Royal Statistical Society, Series B*, **10**, 159-203.

Robert, C. (1992), *L'analyse statistique bayésienne*, Economica, Paris.

Tierney, L. (1990), *LISP-STAT, An object-oriented environment for statistical computing and dynamic graphics*, John Wiley, New York.

Titterington, D.M., Smith, A.F.M., Makov, U.E. (1985), *Statistical analysis of finite mixture distributions*, John Wiley, New York.

Wolfram, S. (1991), *Mathematica: A system for doing mathematics by computer*, Addison-Wesley, Massachusetts

Zhou, X.H., Melfi, C.A., Hui, S.L. (1997), Methods for comparison of cost data, *Annals of Internal Medicine*, **127**, 752-6.