

UNIVERSITÉ DU QUÉBEC

MÉMOIRE PRÉSENTÉ À  
L'UNIVERSITÉ DU QUÉBEC À TROIS-RIVIÈRES

COMME EXIGENCE PARTIELLE  
DE LA MAÎTRISE EN MATHÉMATIQUES ET INFORMATIQUE  
APPLIQUÉES

PAR  
YOUSSEF HANDI

MODÈLE D'AGRÉGATION DES AVIS DES EXPERTS, EN FIABILITÉ  
D'ÉQUIPEMENTS

FÉVRIER 2021

Université du Québec à Trois-Rivières

Service de la bibliothèque

Avertissement

L'auteur de ce mémoire ou de cette thèse a autorisé l'Université du Québec à Trois-Rivières à diffuser, à des fins non lucratives, une copie de son mémoire ou de sa thèse.

Cette diffusion n'entraîne pas une renonciation de la part de l'auteur à ses droits de propriété intellectuelle, incluant le droit d'auteur, sur ce mémoire ou cette thèse. Notamment, la reproduction ou la publication de la totalité ou d'une partie importante de ce mémoire ou de cette thèse requiert son autorisation.

# Table des matières

Table des matières	ii
Liste des tableaux	iv
Table des figures	v
<b>1 Introduction</b>	<b>1</b>
1.1 Problématique . . . . .	2
1.2 Objectif du travail . . . . .	3
1.3 Organisation du mémoire . . . . .	5
<b>2 APPROCHES ET REVUE DE LITTÉRATURE</b>	<b>6</b>
2.1 Notion en fiabilité . . . . .	7
2.1.1 Fiabilité d'un équipement . . . . .	7
2.2 Estimateur bayésien . . . . .	10
2.3 Distributions conjuguées . . . . .	11
2.4 Méthode simulation <i>MCMC</i> . . . . .	12
2.4.1 Chaînes de Markov . . . . .	12
2.4.2 Ergodicité et distribution stationnaire . . . . .	13
2.4.3 Méthodes de Monte-Carlo . . . . .	14
2.4.4 Algorithme <i>MCMC</i> . . . . .	15
2.4.5 L'algorithme de Metropolis-Hastings . . . . .	16
2.5 Modélisation des avis des experts . . . . .	18
2.5.1 Méthode de mélange . . . . .	19

2.5.2	Méthode de la moyenne . . . . .	20
2.6	Qualité d'ajustement d'un modèle bayesien . . . . .	21
2.6.1	Déviance bayésienne . . . . .	21
2.6.2	Analyse prédictif du modèle bayésien . . . . .	22
<b>3</b>	<b>MODÈLE D'AGRÉGATION DES AVIS DES EXPERTS</b>	<b>24</b>
3.1	Calcul de vraisemblance et avis des experts . . . . .	25
3.1.1	Fonction de vraisemblance . . . . .	25
3.1.2	Modélisation des avis des experts . . . . .	26
3.1.3	Calcul de la loi a priori des avis des experts . . . . .	29
3.2	Calcul du DIC et la p-valeur bayésien . . . . .	30
<b>4</b>	<b>ANALYSE ET INTERPRÉTATION DES RÉSULTATS</b>	<b>32</b>
4.1	Comparaison des modèles de combinaison d'avis des experts . . . . .	33
4.2	Validation des résultats d'agrégation des avis des experts . . . . .	36
4.2.1	Validation par les quantiles simulés et théoriques . . . . .	36
4.2.2	Validation par les densités des valeurs simulées du paramètre .	37
4.3	Application du modèle d'agrégation sur les données HQ . . . . .	39
4.3.1	Illustration des données . . . . .	39
4.3.2	Interprétation des résultats . . . . .	39
<b>5</b>	<b>Conclusion</b>	<b>43</b>
	<b>Bibliographie</b>	<b>45</b>
<b>A</b>	<b>Code R :</b>	<b>48</b>

# Liste des tableaux

2.1	<i>Relations entre les différentes fonctions de fiabilité . . . . .</i>	8
2.2	<i>Tableau de quelques distributions usuelles conjuguées . . . . .</i>	11
2.3	<i>Avantage &amp; inconvénient des méthodes d'estimation . . . . .</i>	21
4.1	<i>Durées de bon fonctionnement. . . . .</i>	33
4.2	<i>Avis des deux experts et paramètres des distributions a priori. . . . .</i>	33
4.3	<i>Résultats a posteriori des trois modèles. . . . .</i>	34
4.4	<i>Probabilité a priori pour chaque expert du modèle d'agrégation des avis des experts. . . . .</i>	34
4.5	<i>Avis des experts sur le bon fonctionnement fourni par l'IREQ. . . . .</i>	39
4.6	<i>Taux de bon fonctionnement observé fourni par IREQ. . . . .</i>	39
4.7	<i>Estimation du taux de fiabilité par la méthode d'agrégation. . . . .</i>	40

# Table des figures

1.1	<i>Modèle de prédiction de fiabilités cité dans le travail[1]</i>	4
2.1	<i>Fonction de défaillance en forme de baignoire</i>	9
2.2	<i>Résumé des articles fourni par Hydro Québec</i>	23
4.1	<i>Valeurs simulées du paramètre d'agrégation d'avis des experts</i>	35
4.2	<i>Valeurs simulées du paramètre par la méthode agrégation de la moyenne</i>	35
4.3	<i>Quantile-simulé vs Quantile-théorique</i>	36
4.4	<i>résultat de la régression des quantiles théoriques vs quantile simulé</i>	37
4.5	<i>Densités des valeurs du paramètre générées par les trois méthodes</i>	38
4.6	<i>Valeurs estimées du taux de fiabilité année 2014 par la méthode D'agrégation</i>	40
4.7	<i>Valeurs estimées du taux de fiabilité année 2015 par la méthode D'agrégation</i>	41
4.8	<i>Valeurs estimées du taux de fiabilité année 2016 par la méthode D'agrégation</i>	41
4.9	<i>Valeurs estimées du taux de fiabilité année 2017 par la méthode D'agrégation</i>	42
4.10	<i>Valeurs estimées du taux de fiabilité année 2018 par la méthode D'agrégation</i>	42

# Chapitre 1

## Introduction

Ce projet de recherche a été réalisé en partenariat avec Hydro-Québec (HQ) dans le cadre du programme de la Chaire de recherche HQ en gestion des actifs à l'Université du Québec à Trois-Rivières.

Les actifs des systèmes de transmission à très haute tension sont généralement entretenus et remplacés selon une stratégie basée sur le temps. En raison de la libéralisation des marchés de l'énergie et de la réglementation incitative, les gestionnaires de réseau de transport doivent fonctionner de manière rentable. Économiser de l'argent avec un niveau inférieur de maintenance et de remplacement met en danger la disponibilité des actifs et la sécurité de l'approvisionnement, d'où la nécessité d'utiliser des algorithmes de prédiction et d'optimisation très robuste. En effet, les chercheurs de l'Institut de Recherche d'Hydro-Québec (IREQ) ont développé un projet nommé PRIAD qui inclut plusieurs modules, soit le module d'entrepôt de données, le module de modélisation des actifs, la banque de fiabilité, le simulateur de fiabilité du réseau de transport, le module de risque et le module d'optimisation.

Ce dernier garantira une bonne planification de la gestion des actifs à Hydro-

Québec Transport (HQT). Cela permet d'établir une estimation robuste de la fiabilité du système de transport et d'élaborer une stratégie de maintenance raisonnable.

## 1.1 Problématique

Dans l'analyse de fiabilité d'un équipement en fonctionnement, la vie d'un équipement est souvent modélisée par des lois de probabilité. Ces lois caractérisent le comportement passé de l'équipement afin d'estimer la durée de vie future. Ces tâches sont délicates lors de la présence des données censurées, données non conformes ou manquantes.

Généralement lorsque les données ne sont pas utilisables, nous utilisons des méthodes statistiques afin de les ajuster par estimation des valeurs manquantes en utilisant la méthode de vraisemblance maximale ou bien de l'extrapolation. L'estimation dans le contexte des données censurées se fait par des méthodes non paramétriques basées sur les rangs des durées de vie.

Or, lorsque les données de fiabilité ne sont pas disponibles, l'estimation des paramètres inconnus par les méthodes statistiques est biaisée. Ainsi, les lois statistiques caractérisant le comportement de l'équipement telles que la fiabilité et la défaillance ne sont pas représentatives. Par conséquent, elles ne peuvent pas être utilisées pour optimiser une stratégie de maintenance. Dans ce contexte, nous considérons généralement les lois de défaillance fournies par le fournisseur pour planifier les stratégies de maintenance. Ces lois sont établies sur la base d'essais de laboratoire ou de retour d'expérience du service après-vente, d'où la nécessité d'interroger des experts afin d'avoir une idée sur sa durée de vie future. Donc la question qui se pose est : comment estimer la fiabilité d'un équipement en introduisant les avis des experts dans notre modèle ? Cette question peut se dévisser en quatre sous questions.

1. Comment entraîner un modèle statistique par des avis d'experts sur les paramètres inconnus du modèle ?
2. Comment combiner plusieurs avis d'experts pour identifier un paramètre inconnu ?
3. Comment définir le poids (contribution) de chaque avis d'experts dans l'estimation du paramètre inconnu ?
4. Comment valider les modèles statistiques en tenant compte de plusieurs avis d'experts et l'estimer à partir des données enregistrées ?

Afin de répondre à nos questions, nous allons adopter une méthode basée sur l'inférence Bayésienne. Cette approche permettra de combiner des avis d'experts exprimés sur un ou plusieurs paramètres inconnus avec les données de défaillance déjà enregistrées. Les paramètres inconnus seront ainsi déterminés en considérant les données de défaillance et les avis des experts. Des outils mathématiques seront utilisés afin d'estimer la loi du paramètre proposé par les avis des experts.

## 1.2 Objectif du travail

Après plusieurs rencontres et discussions intéressantes avec l'équipe du projet PRIAD de L'IREQ, nous avons établi l'objectif du présent travail. Il s'agit de déterminer une procédure d'estimation de la fiabilité d'un équipement en fonctionnement en se basant sur les données enregistrées et les avis des experts. L'approche proposée fournit deux méthodes d'estimation des paramètres inconnus. La première se base sur les avis des experts  $\lambda_{predicted}$ , et la deuxième sur les données  $\lambda_{observed}$ , nous utiliserons en suite des méthodes d'optimisation afin de trouver la solution optimale telle qu'expliqué dans le travail de [1], et comme le schéma suivant le résume :

Nos objectifs se résument en cinq points :

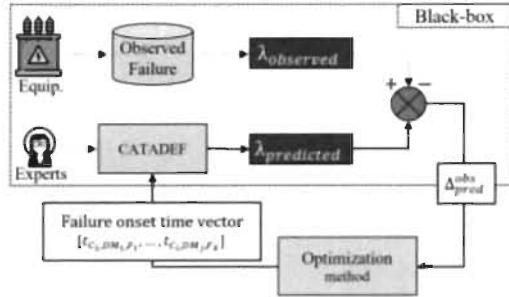


FIGURE 1.1 – Modèle de prédiction de fiabilités cité dans le travail[1]

- Transformer les intervalles des temps de propagation des dégradations données par les experts en des distributions statistiques (estimer les 25 quantiles par une loi normale).
- Identifier la distribution résultante du taux de défaillance prédit par CATA-DEF à travers les distributions des temps de propagation des dégradations par simulations Monte-Carlo ou autres.
- Identifier la méthode afin d’appliquer l’approche bayésienne sur la fusion de la distribution du taux de défaillance observée et prédite.
- Méthode pour l'estimation des pondérations des confiances entre les deux distributions.
- Intégration de la méthode proposée dans le projet PRIAD.

Lors de l’étude des objectifs de l’équipe du projet PRIAD et en utilisant une recherche sur les méthodes déjà utilisées, nous avons proposé un modèle qui permet d’estimer les paramètres inconnus en combinant les avis des experts. Les objectifs projetés sont :

- (a) Développer un modèle statistique d’agrégation des avis d’expert.
- (b) Choisir un modèle d’agrégation des avis des experts basé sur le critère de sélection DIC (Deviance information criterion).
- (c) Choisir un modèle statistique a posteriori sur les paramètres d’intérêt en utilisant les avis d’experts retenus par le critère choisi en (b).

Après l'estimation de la durée de vie d'un équipement par les avis des experts,

nous comparerons la loi obtenue à celle déterminée à partir des données enregistrées. Nous aurons aussi ensuite recours aux tests d'ajustement afin de trouver la loi qui décrit mieux la durée de vie.

### 1.3 Organisation du mémoire

Le présent mémoire se décompose de quatre chapitres. Le deuxième chapitre présente une introduction sur l'inférence Bayésienne ainsi que quelques propriétés sur les estimateurs. Nous discuterons ensuite des certaines méthodes de simulations basées sur les chaines de Markov appelées MCMC (Monte-Carlo par Chaine de Markov). Ces dernières permettent d'ajuster l'estimation des paramètres inconnus par la méthode Bayésienne proposée. Le troisième chapitre traite du modèle d'agrégation des avis d'experts afin de bien estimer la loi a priori qui combine les avis des experts. Le dernier chapitre présentera une simulation des données de bon fonctionnement des machines ainsi, qu'une discutions sur les résultats trouvés. Le cinquième chapitre présentera la conclusion du travail et propose d'autres perspectives afin de bien améliorer les méthodes traitées.

# Chapitre 2

## APPROCHES ET REVUE DE LITTÉRATURE

Ce chapitre est inspirer des notes du cours [2] et [3] .

Dans ce chapitre, nous introduirons quelques approches utilisées en estimation de la fiabilité et combinaison des avis des experts sur un paramètre inconnu. Nous présenterons la méthode d'estimation bayésienne ainsi que quelques méthodes utilisées afin de bien estimier la loi a priori, qui combine les avis des experts. Une revue de littérature sur ces méthodes sera aussi présentée. Par la suite nous allons introduire des critères statistiques permettant de mesurer la qualité d'estimation ainsi qu'une revue sur les méthodes de combinaison des avis des experts. Nous terminerons par présenter une grille qui résume les articles sur lesquels nous nous sommes basés pour déterminer la méthode utiliser dans le chapitre 3.

## 2.1 Notion en fiabilité

### 2.1.1 Fiabilité d'un équipement

Selon la norme *AFNOR X06501*, la fiabilité est l'aptitude d'un système à accomplir une fonction requise dans des conditions d'utilisation et pour une période de temps bien déterminées . En pratique, la fiabilité est directement liée à la durée de vie d'un équipement. Cette durée de vie est une variable aléatoire, notée  $T$  et caractérisant le passage aléatoire d'un équipement d'un état de fonctionnement à un état de défaillance selon une loi de probabilité qui peut être connue ou inconnue. La durée de vie  $T$  s'exprime en unités de temps, en nombre de cycles ou en unités produites par l'équipement.

Le modèle permettant de calculer la probabilité de survie d'un équipement est appelé « fonction de fiabilité ». Elle s'exprime en fonction de la durée  $t$ , une réalisation de la variable aléatoire  $T$  est donnée par  $\mathcal{R}_T(t) = \mathbb{P}(T > t)$ . On la nomme la probabilité de survie. La probabilité que l'équipement tombe en défaillance avant le temps  $t$  est donné par le risque de défaillance, noté par :

$$F(t) = 1 - \mathcal{R}_T(t) = \mathbb{P}(T \leq t).$$

La fonction de densité de défaillance  $f$ , permet d'obtenir la probabilité instantanée de défaillance d'un équipement entre les instants  $t$  et  $t + dt$

$$\int_t^{t+dt} f(t) dt = \mathbb{P}(t < T < t + dt),$$

et la fonction de fiabilité est donnée par

$$\mathcal{R}_T(t) = \int_t^{+\infty} f(v) dv.$$

Bien que la fonction  $\mathcal{R}_T(t)$  peut s'écrire aussi en fonction du taux de défaillance  $\lambda_T(t)$ .

Ce taux calcule la probabilité qu'un équipement tombe en défaillance par unité de temps. En d'autres mots, c'est la vitesse d'arrivée des défaillances par unité de temps. La fonction  $\lambda_T(t)$  donne la probabilité de défaillance d'un équipement entre  $t$  et  $t + dt$  sachant que l'équipement a survécu jusqu'à l'instant  $t$ . Le tableau 2.1 résume les différentes relations entre les fonctions de fiabilité.

Fonction de fiabilité	$f_T(t)$	$\mathcal{R}_T(t)$	$F_T(t)$	$\lambda_{T(t)}$
$f_T(t)$	-	$\int_t^{+\infty} f(v)dv$	$\int_0^t f(v)dv$	$f(t)/\int_t^{+\infty} f(v)dv$
$\mathcal{R}(t)$	$-d\mathcal{R}(t)/dt$	-	$1 - \mathcal{R}(t)$	$(-d\mathcal{R}(t)dt)/\mathcal{R}(t)$
$F(t)$	$dF(t)/dt$	$1 - F(t)$	-	$(dF(t)dt)/F(t)$
$\lambda(t)$	$\lambda(t) e^{-\int_0^t \lambda(v)dv}$	$e^{\int_0^t \lambda(v)dv}$	$1 - e^{-\int_0^t \lambda(v)dv}$	-

TABLE 2.1 – Relations entre les différentes fonctions de fiabilité

Habituellement, il existe deux méthodes pour estimer la fonction de défaillance. Il s'agit des méthodes non paramétriques et paramétriques.

(a) **Méthode d'estimation non paramétrique** : cette méthode se base sur les rangs de la défaillance dans un échantillon donné. Les méthodes les plus utilisées sont les méthodes de Kaplan-Meier [4] et de Johnson [5] sont les plus utilisées en fiabilité. Ces méthodes donnent une estimation de la probabilité de défaillance d'un équipement dans une durée du temps définie. De plus, il n'y a aucune exigence sur la distribution de la population soit caractérisée par certains paramètres. Les méthodes non paramétriques sont utiles lorsque l'hypothèse de normalité ne tient pas et que l'effectif d'échantillon est faible. Cela dit, dans les tests non paramétriques, nos données reposent également sur des hypothèses. Par exemple, il est essentiel de supposer que les observations des échantillons sont indépendantes et obéissent à la même loi. De même, dans les plans à deux échantillons, il est nécessaire de supposer que la forme et la dispersion des lois sont similaires.

(b) **Méthode d'estimation paramétrique** : contrairement aux méthodes non paramétriques, les méthodes paramétriques prennent comme hypothèse que les

données des durées de vie suivent une loi de probabilité connue. Cette dernière est caractérisée par un ou plusieurs paramètres. Les données observées sont utilisées pour calculer des estimations des paramètres ciblés. Parmi les méthodes paramétriques utilisées en fiabilité sont les modèles Exponentiel et Weibull. Comme mentionné dans le livre [6], dans ces approches le taux de défaillance est considéré constant ce qui implique que l'occurrence des défaillances suit un processus de Poisson homogène de paramètre  $\lambda$  ( $\lambda = cst$ ). Cette approche est souvent utilisée dans des équipements électroniques et dans la défaillance se réalise soudainement et indépendamment de son âge et elle est aussi utilisée dans les équipements multicomposants [7].

Le second modèle est celui de Weibull, qui est aussi un modèle exponentiel, avec un paramètre de plus qui inclut la flexibilité du modèle à modéliser des modes de défaillance plus complexe (usure, fatigue, etc.) telle expliqués dans le livre [8]. Le taux de défaillance du modèle de Weibull est bien adapté pour décrire les trois phases de vie d'un équipement suivant son paramètre de forme  $\beta$  :

- $\beta < 1$  : Phase défaillances de jeunesse, on a  $\lambda_T(t)$  décroît avec le temps.
- $\beta = 1$  : Phase de la vie utile, on a  $\lambda_T(t)$  est constant.
- $\beta > 1$  : Phase usure,  $\lambda_T(t)$  augmente avec le temps.

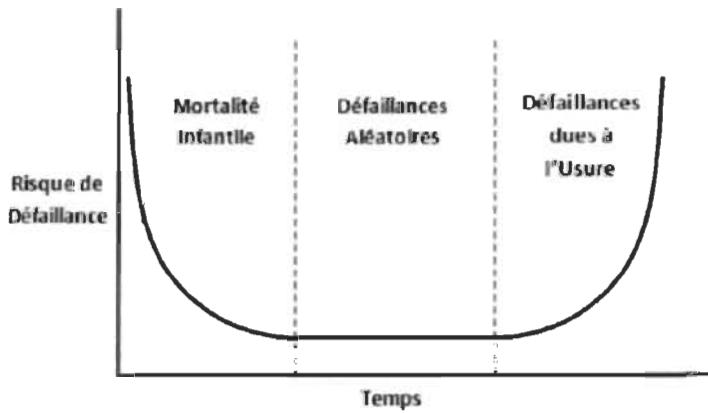


FIGURE 2.1 – Fonction de défaillance en forme de baignoire

Pour un modèle de fiabilité donné, il est possible de calculer sa durée de vie

moyenne comme étant l'espérance mathématique de la variable aléatoire T :

$$\mathbb{E}[T] = \int_0^{+\infty} \mathcal{R}_T(t) dt = MTBF$$

*MTBF* désigne *Mean Time Before Failure*. Pour les systèmes réparables.

## 2.2 Estimateur bayésien

Contrairement à la méthode de vraisemblance, la méthode d'estimation bayésienne considère que les paramètres sont des variables aléatoires. Cette méthode utilise la règle de Bayes dans le calcul des probabilités conditionnelles entre évènements. Chaque paramètre inconnu sera modélisé par une densité de probabilité  $\Pi(\theta)$  appelée distribution a priori en plus dans l'estimateur de Bayes va tenir compte des observations dans la fonction a posteriori sur le paramètre obtenu en combinant la fonction de vraisemblance caractérisée par les observations  $\Pi(x|\theta)$  et la fonction a priori qui contient l'information sur le paramètre. La fonction a posteriori est donnée par le théorème de Bayes et se calcule de la manière suivante suivant :

$$\Pi(\theta|x) = \frac{\Pi(x|\theta) \Pi(\theta)}{\int \Pi(x|\theta) \Pi(\theta) dx}.$$

Donc l'estimateur de Bayes est tout simplement, la moyenne de la loi a posteriori résultante définie dans la formule précédente. L'estimateur bayésien est défini par la moyenne de la loi  $\Pi(\theta|x)$ .

$$\hat{\theta} = \mathbb{E}(\theta|x).$$

On peut constater que l'information a priori sur le paramètre sera actualisée directement à partir des données observées via la fonction de vraisemblance. D'après les articles [9] et [1] fournis par l'IREQ sur l'estimation et l'optimisation du paramètre de

fiabilité du système de transport électrique chez Hydro-Québec, cette méthode nous aidera beaucoup afin de bien estimer notre paramètre d'intérêt sans utiliser deux estimations à partir des données enregistrées et les avis des experts questionnés sur la fiabilité de l'équipement. Un des avantages de cette méthode est si la fonction de vraisemblance  $\Pi(x|\theta)$  est conjuguée avec distribution a priori  $\pi(\theta)$ , dans ce cas la distribution a posteriori sera de même forme que celle de la distribution a priori.

## 2.3 Distributions conjuguées

**Définition 2.3.1.** Soit  $\mathcal{F}$  une famille de densités  $f(x|\theta)$  indexées par  $\theta$ . Une famille  $\Pi$  de densités a priori sur  $\theta$  est dite conjuguée de  $\mathcal{F}$  si la densité a posteriori fait partie de la famille  $\Pi$  pour tout  $f \in \mathcal{F}$ , tout  $\pi \in \Pi$ .

Souvent la forme analytique de la distribution a posteriori est difficile de la calculer. Des méthodes de simulation ont été proposées afin de trouver une forme analytique de la fonction résultante  $\Pi(\theta)$ . Le but principal de ces méthodes est de générer un échantillon suivant la loi résultante  $\pi(\theta|x)$ . Les algorithmes les plus utilisés dans la pratique sont de type simulation Monte-Carlo par chaîne de markov (MCMC). L'algorithme de simulation MCMC le plus utilisé est celui de Metropolis Hastin [10]. Le fonctionnement de ces algorithmes sera détailler dans la section suivante.

$\Pi(x \theta)$	$\Pi(\theta)$	$\Pi(\theta x)$
$Normale(\theta, \sigma^2)$	$Normale(\mu, \tau^2)$	$Normale(\tau(\sigma^2\mu + \tau^2x), \tau\sigma^2\tau^2)$ avec $\tau^{-1} = \sigma^2 - \tau^2$
$Poisson(\theta)$	$Gamma(\alpha, \beta)$	$Gamma(\alpha + x, \beta + 1)$
$Gamma(\alpha, \beta)$	$Gamma(\tau, \theta)$	$Gamma(\alpha + \tau, \beta - x)$

TABLE 2.2 – Tableau de quelques distributions usuelles conjuguées

## 2.4 Méthode simulation MCMC

Les méthodes de Monte-Carlo par chaînes de Markov (MCMC) forment une classe d’algorithmes permettant de simuler des variables aléatoires provenant de distributions complexes, d’abord utilisées en physique statistique puis développées subséquemment de façon plus générale. Elles sont utilisées en applications de la statistique bayésienne, dans des domaines aussi variés que l’analyse du mouvement d’une molécule dans l’espace. Les méthodes de Monte-Carlo sont nées à Los Alamos, Nouveau-Mexique pendant la Seconde Guerre mondiale, ce qui a finalement abouti à l’algorithme Metropolis au début des années 1950. Alors que les méthodes de Monte-Carlo étaient utilisées par ce temps, MCMC a été rapproché de l’aspect pratique de la statistique par les travaux de Hastings dans les années 1970. Elles sont relativement simples à implémenter et ne requièrent souvent que la connaissance de la fonction de densité cible à une constante près, ce qui les rend intéressantes dans de nombreuses situations. Cependant, une implémentation naïve peut mener à des temps de calcul très longs, puisque la convergence de ces méthodes est relativement lente lorsqu’elles ne sont pas bien calibrées à une situation donnée.

### 2.4.1 Chaînes de Markov

Dans cette section, nous allons citer la définition d’une Chaîne de Markov et le comportement à long terme, ainsi quelque notion théorique des méthodes Monte-Carlo.

**Définition 2.4.1.** *On appelle chaîne de Markov à temps discret toute suite de variables aléatoires  $X_t$  à valeurs dans un espace d’états dénombrable telle que :*

$$\mathbb{P}(X_t | X_{t-1} \dots X_0) = \mathbb{P}(X_t | X_{t-1}),$$

est la propriété markovienne.

La probabilité cette suite à un temps donné  $t$ , sachant états précédentes, ne dépend en fait que de la valeur du temps précédente. On étudiera ici uniquement les chaînes dites homogènes où les probabilités  $\mathbb{P}(X_t | X_{t-1})$  sont indépendantes de la valeur  $t$ . Donc une chaîne de Markov homogène est complètement définie par la valeur ou la distribution de  $X_0$  et les probabilités de transition en un pas

$$p_{ij} = \mathbb{P}(X_{t+1} = j | X_t = i).$$

## 2.4.2 Ergodicité et distribution stationnaire

On dit qu'une distribution est stationnaire, lorsqu'elle existe, est une distribution de probabilités sur un support  $S$  et notée par le vecteur  $\pi = (\pi_i)_{i \in S}$  où  $\pi_i = \mathbb{P}(Z = i)$ ,  $Z$  est une distribution de la loi  $\pi$ , et il respecte l'équation suivante  $\pi = \pi \times P$ . En d'autres termes, on a

$$\forall j : \pi_j = \sum_i \pi_i P_{ij}.$$

Dans tel cas, la distribution stationnaire existera toujours et sera unique. De plus, on aura alors, pour tous les états  $i, j \in S$  :

$$\lim_{n \rightarrow +\infty} P_{ij}^{(n)} = \pi_i.$$

Il s'agit du théorème ergodique pour les chaînes de Markov homogènes sur un espace discret. En effet, une chaîne de Markov ergodique convergera en loi vers sa distribution stationnaire, peu importe la distribution de la variable initiale  $\{X_0\}$ .

### 2.4.3 Méthodes de Monte-Carlo

Les méthodes de Monte-Carlo sont définies comme toute technique de calcul utilisant des procédés aléatoires ou pseudo-aléatoires. Dans un contexte statistique, on s'intéresse à un sous-ensemble de ces méthodes, soit les techniques d'échantillonnage aléatoire numérique visant à calculer des intégrales en plusieurs dimensions. Le problème fondamental considéré est le suivant. Soit  $X$  une variable aléatoire de support  $S \subset \mathbb{R}^d$  et de densité  $f$ , et  $h$  une fonction quelconque définie sur  $S$ . On cherche à calculer une bonne approximation de :

$$I = \mathbb{E}_f[h(X)] = \int_S h(x) f(x) dx.$$

Afin de bien calculer la quantité  $I$ , on génère un certain nombre  $n$  de réalisations  $x_i \in S$  avec  $i = 1, \dots, n$  indépendants et identiquement distribués (i.i.d.) selon la distribution de  $X$ , puis on estime  $I$  par

$$\bar{h}_n = \frac{1}{n} \sum_{i=1}^n h(x_i).$$

**Propriétés 2.4.1.** (a) *Par la loi des grands nombres on a  $\bar{h}_n$  converge en loi vers  $\mathbb{E}_f[h(X)]$ .*

(b) *La quantité doit être finie  $\mathbb{E}_f[h(X)^2]$ .*

(c) *Sous la condition (b)  $\frac{\bar{h}_n - I}{\sqrt{\text{Var}(h(X))}}$  converge vers la loi Normale(0,1).*

Ces techniques sont surtout utiles dans les problèmes en grande dimension, où les méthodes numériques traditionnelles perdent de leur efficacité. La difficulté dans l'application de ces méthodes est de trouver une façon de générer efficacement un échantillon de variables i.i.d. de densité  $f$ .

#### 2.4.4 Algorithme MCMC

Les méthodes de Monte-Carlo par chaînes de Markov permettent d’élargir l’échelle des loi de probabilités. cet algorithme permet de simulées numériquement telle expliquée dans le livre [11]. Elles sont relativement simples à implémenter et n’exiger souvent que la connaissance de la fonction de densité cible, ce qui les rend intéressantes dans de nombreuses situations. Par contre, un choix de la loi cible naïve peut aboutir à des temps de calcul très longs, puisque la convergence de ces méthodes est relativement lente lorsqu’elles ne sont pas bien calibrées à une situation donnée.

Nous représenterons en premier l’angle théoriques de cette méthode ensuite nous introduisant l’algorithme de Metropolis-Hastings et ses propriétés. L’idée principale des MCMC est de simuler des observations d’une distribution de densité  $\pi$  donnée en exploitant une chaîne de Markov ergodique  $\{X_t\}$  dont la distribution stationnaire est  $\pi$ . Le théorème ergodique garantit la convergence en loi de  $X_t$  vers une variable aléatoire de densité  $f$  et par conséquent, quelque soit la valeur initiale  $X_0$  :

$$\frac{1}{n} \sum_{i=1}^n h(x_i) \rightarrow \mathbb{E}_f[h(X)].$$

L’algorithme MCMC toute méthode permettant de concevoir des variables d’une distribution en utilisant une chaîne de Markov ergodique ayant celle-ci comme distribution stationnaire. afin de construire un tel processus, il faut donc déterminer une probabilité de transition  $P(x, .)$ , de telle façon à respecter les conditions du théorème ergodique et choisir la bonne distribution stationnaire. Afin de bien adopter cette distribution, nous appliquerons le résultat suivant.

**Proposition 2.4.1.** *Soit une chaîne de Markov ayant un probabilité de transition  $P(x, .)$  et une distribution  $\pi(.)$  définis sur le même espace d’états  $S$ . Si  $P$  accaparer*

*la propriété de réversibilité par rapport à  $\pi$ , c'est-à-dire si, pour tout  $x, y \in S$ , on a*

$$\pi(x) P(x, y) = \pi(y) P(y, x),$$

*alors la distribution stationnaire de la chaîne est  $\pi$ .*

**Démonstration 2.4.1.** *On a la stationnarité si, pour tout  $y \in S$  :*

$$\int_S \pi(x) P(x, y) dx = \pi(y).$$

*Or sous l'hypothèse de réversibilité, on a pour tout  $y$*

$$\int_S \pi(x) P(x, y) dx = \int_S \pi(y) P(y, x) dx = \pi(y) \int_S P(x, y) dx = \pi(y).$$

La réversibilité est simple à revivifier que l'ergodicité. En effet, cette notion permet de construire des probabilités de transition adapté pour une distribution stationnaire attendus.

## 2.4.5 L'algorithme de Metropolis-Hastings

La simplicité de l'algorithme de Metropolis-Hastings [12], pour une densité cible  $\pi(\cdot)$  donnée, nécessite seulement une valeur initial  $X_0$  et une distribution conditionnelle de densité telle que  $q(x, y) := q(y|x)$ . Cette dernière servira à proposer un candidat l'état initial de la chaîne  $X_t$ , et un autre candidat pour l'état suivant de la chaîne. Ce candidat sera ensuite accepté ou rejeté avec une probabilité, choisie de manière à ce que la chaîne soit réversible par rapport à  $\pi(\cdot)$ . Précisément, à une transition donnée  $t \in \mathbb{N}$ , les étapes suivantes sont effectuées :

- (1) Depuis la valeur  $X_t = x$ , on génère  $Y_{t+1} = y$  selon la distribution de densité

$$q(y|x).$$

(2) On définit

$$X_t = \begin{cases} Y_{t+1} & \text{avec probabilité } \alpha(x, y) \\ X_t & \text{avec probabilité } 1 - \alpha(x, y), \end{cases}$$

où la probabilité  $\alpha(x, y)$  de la forme générale

$$\alpha(x, y) = \frac{s(x, y)}{1 + r(x, y)},$$

d'où  $r(x, y)$  est le ratio

$$\frac{\pi(x) q(x, y)}{\pi(y) q(y, x)},$$

et la fonction  $s$  est définie de façon à ce que

$$s(x, y) = s(y, x) \pi(y) q(y, x).$$

ce qui nous donne,

$$\alpha(x, y) = \min \left( 1, \frac{\pi(x) q(x, y)}{\pi(y) q(y, x)} \right),$$

qui correspond au choix

$$s(x, y) = \min (1 + r(x, y), 1 + r(y, x)).$$

(3) Si la densité  $q$  est symétrique  $q(y|x) = q(x|y)$ , le rapport devient tout simplement

$$\alpha(x, y) = \min \left( 1, \frac{\pi(x)}{\pi(y)} \right).$$

**Remarque sur le fonctionnement de l'algorithme MH :** Le choix de la distribution  $q(x, y)$  qui influencera la performance de l'algorithme. Un choix optimale de cette distribution favorisera une exploration rapide de l'espace des valeurs générées et accélérera la convergence de la chaîne vers la distribution stationnaire. Le but prin-

cipale est d'avoir une variabilité responsable des candidats possibles, de façon à tenir en compte les fluctuations tout en restreignant la probabilité de rejet  $\alpha(x, y)$ . Dans la pratique, on éliminera les échantillons obtenus en premières étapes de l'algorithme, qui sont influencées par le choix de la valeur initiale. De même, afin d'éviter qu'il y ait de corrélation entre des valeurs de l'échantillon, on pourrait choisir de conserver uniquement une valeur sur 10. Puisque les probabilités  $\alpha(x, y)$  ne dépendent de la distribution cible que sous la forme du rapport  $\pi(x)/\pi(y)$ , il est souvent possible d'appliquer l'algorithme sans connaître la constante de normalisation de la densité de  $\pi$ .

## 2.5 Modélisation des avis des experts

Dans cette section, nous allons introduire les méthodes permettant de combiner plusieurs avis des experts exprimés sur un paramètre inconnu. Tel que vu dans la section précédente, une distribution a priori sur un paramètre inconnu combinée avec la distribution de vraisemblance des données observées, qui évalue la distribution a posteriori de ce paramètre par l'inférence bayésienne. La distribution a priori est généralement construite sur la base de connaissances sur le paramètre d'intérêt par la personne qui va prendre la décision ou pour prédire le futur. Dans la littérature, il existe deux méthodes pour combiner les avis des experts : la combinaison mathématique des distributions de probabilités, et la combinaison se basant sur un accord ou consensus entre les experts [13]. Il existe deux méthodes utilisées dans la littérature afin de combiner les avis des experts. La méthode qu'on va utiliser dans ce travail consiste à combiner plusieurs distributions de probabilités et la combinaison se basant sur l'accord entre tous les avis reçus. La méthode est simple, il repose sur un questionnaire qui doit être donné à chaque expert afin de déduire une distribution a priori résultante comme un accord entre les différentes réponses fournies par les experts. Cependant s'il y a un désaccord entre les experts la distribution ne peut pas

être défini car ceux-ci n'arrivent pas à trouver une fonction commune. Après cette brève introduction, nous présenterons les deux principales méthodes.

### 2.5.1 Méthode de mélange

Supposons que nous avons  $N$  experts  $(E_1, \dots, E_N)$  qui ont été consultés et que nous avons associé à chacun une distribution a priori  $\Pi_i(\theta)$  sur un paramètre inconnu  $\theta$  (ou un vecteur de paramètres). Le problème est de déterminer à partir des  $\Pi_i(\theta)$  la distribution a priori unique  $\Pi(\theta)$  qui décrit ainsi un consensus entre les avis des experts. La méthode du mélange est une approche principale dans les travaux de [14] et [15], une combinaison linéaire des distributions a priori de chaque expert :

$$\Pi(\theta) = \sum_{j=1}^N \omega_j \Pi_j(\theta), \quad \Pi_j(\theta) > 0, \quad \sum_{j=1}^N \omega_j = 1.$$

Une restriction de cette méthode est que les poids  $\omega_i$  doivent être fixés et connus d'avance avant de procéder au mélange des avis. Leur somme doit être égale à 1. Le seul problème avec cette méthode est comment estimer les poids de chaque expert dans le mélange. En plus est-ce possible de donner un poids très important à l'expert  $i$  si il est jugé meilleur ou plus performant que les autres. Plusieurs travaux ont traité la question d'estimation du poids. Nous allons donner quelque règle afin d'estimer les poids de chaque expert. Nous allons commencer par attribuer des poids égaux à chaque expert dans ce cas on présume qu'il y a une grande différence entre les experts. D'autres approches consistent à donner un poids à chaque expert proportionnel à un ordonnancement des experts. Ordonner les experts suivant leur performance attribuer un poids  $r / \sum_{r=1}^N r$ , où  $r = 1, 2, \dots, N$ . Estime le poids sellant la performance passée.

L'avantage de cette méthode permet de prendre en compte les valeurs possibles par les experts, mais il peut influencer notre estimateur, car l'erreur va être augmentée.

### 2.5.2 Méthode de la moyenne

La méthode de la moyenne a été le but de plusieurs travaux dans la littérature [16]. L'avantage de cette méthode est qu'elle est simple à implémenter, mais son inconvénient est qu'elle peut ne pas tenir compte des incertitudes que les experts prévoient sur les paramètres inconnus. Cette méthode constitue un accord entre les différents experts qui sont consultés. Cette méthode nécessite aussi l'utilisation d'un même type de distribution a priori pour tous les experts consultés. Pour simplifier le développement de la méthode, nous allons supposer que la distribution a priori pour chaque expert va être représentée par une loi bêta( $\mu_i, \beta_i$ ). La distribution a priori résultante de la combinaison avec la méthode de la moyenne sera une loi bêta( $\mu, \beta$ ) avec  $\mu = \sum_{j=1}^N \mu_j/n$  et  $\beta = \sum_{j=1}^N \beta_j/n$ .

La méthode de la moyenne, est simple à implémenter, est plus avantageuse à utiliser si les avis des experts sont homogènes. Cependant, l'utilisation de la méthode du mélange est recommandée si les avis des experts sont différents, car dans ce cas, la méthode de la moyenne ne permet pas de tenir compte de l'incertitude des experts puisqu'elle considère une distribution a priori résultante qui est une référence pour tous les experts. Dans notre travail nous allons utiliser la méthode du mélange, car il donne des résultats très convaincants. Ces dernières vont être présentées dans le prochain chapitre. En particulier, nous allons tester les deux méthodes sur deux avis experts l'un est qualifié bon et l'autre mauvais. Le tableau suivant résume les avantages et les inconvénients de chaque méthode

Méthode	Avantage	Inconvénient
Méthode de mélange	-Permet d'inclure toutes les valeurs possibles -Inclure l'erreur d'estimation	-calcule très complexe
Méthode de la moyenne	Simple à implémenter	-Mauvaise estimation de la loi apriori si les experts n'ont pas le même champ d'expertise -ne permet pas de tenir compte de l'incertitude

TABLE 2.3 – Avantage &amp; inconvénient des méthodes d'estimation

## 2.6 Qualité d'ajustement d'un modèle bayésien

Dans cette section, nous allons introduire des quantités que nous allons utiliser afin de choisir le bon modèle lors de notre modélisation. Comme dans la statistique fréquentiste, la qualité d'ajustement d'un modèle donné peut être évaluée à partir de plusieurs techniques. Parmi ces techniques la déviance bayésienne et l'analyse du caractère prédictive du modèle sont les plus utilisés.

### 2.6.1 Déviance bayésienne

Le Deviance Information Criterion (DIC) telle mentionnée dans le travail [17] s'appuie sur la déviance, qui s'écrit comme suit :

$$DIC = \bar{D} + P_D,$$

avec

$$\bar{D} = -2 \log (\pi(\theta|x)).$$

L'espérance a posteriori de la déviance qui peut être considérée comme une mesure d'ajustement aux données représentées par  $\pi(x|\theta)$  de la vraisemblance et  $P_D = \bar{D} - D(\bar{\theta})$  est appelé le nombre effectif des paramètres.  $D(\bar{\theta}) = -2 \log(\pi(\bar{\theta}|x))$ , avec  $\bar{\theta}$  est la moyenne a posteriori des paramètres  $\theta$ .  $P_D$  est considéré comme un facteur mesurant la complexité du modèle. En pratique, le calcul du DIC se fait avec les méthodes d'échantillonnage MCMC (algorithme de Metropolis-Hastings). À partir d'un échantillon des valeurs obtenu  $\theta$ , il est possible de le calculer admettant qu'on a observé  $L$  mesure de  $\theta$ .

$$\bar{D} = \frac{1}{L} \times (-2 \log \left( \prod_{k=1}^L (x|\theta^{(k)}) \right)).$$

La sélection du modèle sera gérée par le DIC le plus petit possible.

## 2.6.2 Analyse prédictif du modèle bayésien

L'analyse du modèle bayésien est une technique utilisée dans la mesure de la qualité d'ajustement du modèle. Simple à utiliser et à implémentez facile à interpréter [18]. Le principe est de simuler un échantillon  $x^{sim}$  à partir de la distribution à posteriori de  $\theta$  et la comparer avec les données observées  $x$ , si l'échantillon simulé est plus semblable aux données observées alors on conclut que la qualité d'ajustement du modèle est bonne. Pour la statistique utiliser afin de comparer les modèles est la valeur-p Bayésienne. Cette probabilité est définie comme étant la probabilité que l'échantillon simuler soit supérieur ou égale à l'échantillon observé. Règle de décision est donc, une valeur-p proche de 0 ou 1, indique que le modèle ne représente pas ou ne prédit pas les données observées, et une valeur-p proche de 0.5, indique que le modèle prédit bien les données observées.

Pour conclure ce chapitre, nous allons résumer tous les article [19], [9], [20] [1] fourni par l'équipe PRIAD

#	Auteurs/Titres/Journal/Année	Carréristiques	Objectif	Méthodologie
1	- ALBERT, I. DONNET, S. - Combining experts opinions in prior elicitation. - 2012	- anglais	Combiner les avis des expert de different domaine de compétence on une seule fonction de densité	- utilisé la même pondération pour les expert - utilisé des méthodes d'optimisation afin de trouvé le bon paramètre de la loi a priori
2	- Marco Laumanns, Jiri osenask - Bayesian Optimization Algorithms - ScienceDirect - 2010	- anglais - révisé comité scientifique	Examiner l'utilité du concepte de la degendance dans l'optimisation multi-objectif afin de bien estimer la distribution a priori optimal.	- Introduire les réseaux bayesien, afin d'estimer la probabilité conjointe de l'échantillon observé.
3	- Blancke, Olivier and Tahan, Antoine and Komljenovic, Dragan and al - Application Bayesian Statistical Decision Theory for maintenance - elsevier - 2018	- anglais - révisé comité scientifique	- Introduction de l'approche Bayesienne de la combinaison du risque de défaillance et des conséquences économiques.	- Introduire un modèle d'agrégation des avis d'experts - Développer des méthodes d'arbres de décision pour objectif d'optimiser et la prévention de la maintenance.
4	- Alain,Coté Olivier, Blancket, and al - Combining historical data and domain expert knowledge using optimization to model electrical equipment reliability - 2020	- anglais - Révisé comité scientifique	- Proposition d'une approche hybride en combinant les données de défaillance ainsi les avis des expert est trouvé une solution optimal.	- Introduire une approche hybride se basant sur les avis des experts et les données - Introduire une approche se base sur la physique des équipements même si les données historique sont rare - Pésentation de l'algorithme qui combine différente sources d'information.

FIGURE 2.2 – Résumé des articles fournis par Hydro Québec

## Chapitre 3

# MODÈLE D'AGRÉGATION DES AVIS DES EXPERTS

Dans cette section, nous allons expliquer le modèle d'agrégation des avis des experts exprimés sur un paramètre inconnu. Nous débuterons par présenter une introduction et les démarches à suivre afin de montrer l'utilité de cette méthode. Dans ce travail, nous considérons un équipement neuf mis en exploitation au temps initial  $t = 0$ . Nous supposons que les durées de bon fonctionnement de cet équipement suivent une loi exponentielle de paramètre  $\lambda$ . le fournisseur fournit un taux de défaillance de base noté  $\lambda_0$ .

L'équipement sera utilisé pendant plusieurs périodes et les périodes de bon fonctionnement ainsi que le nombre de défaillances seront enregistrées sur chaque période de temps où, une vraisemblance sera calculée. À l'instant  $t = 0$  chaque expert sera consulté afin de donner leur avis sur le taux de défaillance de la période subséquente. Les experts se basent sur le comportement de l'équipement au cours de la période précédente ou qu'il subira à une maintenance, inspection ou bien une réparation. Au cours de la période ultérieure pour donner leurs avis, et aussi une distribution a priori

sur le taux de défaillance au début de chaque période. La vraisemblance de chaque période sera modélisée par un processus de poisson homogène de paramètre  $\lambda$ . Quand un avis commun pour tous les experts est fourni et que la vraisemblance soit connue nous serons capables d'effectuer une inférence bayésienne sur le paramètre  $\lambda$ .

Nous aurons donc une distribution a posteriori sur le paramètre  $\lambda$  caractérisé par sa moyenne et son écart-type. Nous aurons également un classement des avis des experts qui ont été formulés au début de la période. Ce classement sera validé à l'aide de trois tests statistiques déjà mentionné : DIC, la valeur-p et des tests non paramétriques afin de comparer la tendance des données avec les avis des experts sur plusieurs périodes.

## 3.1 Calcul de vraisemblance et avis des experts

Dans un premier temps nous allons construire notre fonction de vraisemblance à partir des données de bon fonctionnement de l'équipement. Ensuite, nous allons modéliser la distribution a priori sur le paramètre  $\lambda$  pour chaque expert en fonction de son avis. Le modèle d'agrégation des avis des experts sera ensuite formulé à l'aide de l'inférence bayésienne.

### 3.1.1 Fonction de vraisemblance

Prenant une période  $k$  avec  $k \in \{1, 2, \dots, K\}$ , combinée avec les données de bon fonctionnement de l'équipement que nous avons  $x_{i,k}$  avec  $i \in \{1, \dots, n_k\}$ , où pour chaque période  $k$  on a  $n_k$  défaillance enregistrée. Aussi, les durées de bon fonctionnement de l'équipement suivent une loi  $\exp\{\lambda\}$  avec sa densité  $f(x, \lambda) = \lambda e^{-\lambda x}$ . La fonction de vraisemblance est donc donnée par

$$L_k(X, \lambda) = \prod_{i=1}^{n_k} f(x_{i,k}; \theta) = \prod_{i=1}^{n_k} \lambda e^{-\lambda x_{i,k}} = \lambda^{n_k} e^{-\lambda \sum_{i=1}^{n_k} x_{i,k}}, \quad (3.1)$$

où  $X$  désigne l'ensemble de toutes les durées de bon fonctionnement de l'équipement pour la période  $k$ .

Cette vraisemblance sera combinée avec la distribution a priori sur  $\lambda$  afin d'appliquer l'inférence bayésienne. Pour le paramétré  $\lambda$ , on va l'estimer par l'estimateur de maximum de vraisemblance d'une loi exponentielle  $\hat{\lambda} = n_k / \sum_{i=1}^{n_k} x_{i,k}$ .

### 3.1.2 Modélisation des avis des experts

Dans cette section, nous allons supposer qu'au début de chaque période, plusieurs experts en fiabilité donnent leurs avis à propos du taux de fiabilité de l'équipement  $\lambda$ . Les experts peuvent être questionnés sur la moyenne et l'écart-type du  $\lambda$ . Dans notre cas, on va chercher les bornes minimale et maximale du paramètre  $\lambda$ . Supposant que nous avons  $N$  experts et chaque expert  $i$  donne une estimation sur les bornes minimale et maximale du paramètre  $\lambda$ . Donc on va y avoir les quantités suivantes  $\lambda_{\min,i}$  et  $\lambda_{\max,i}$ . À partir des deux informations de chaque expert  $i$  nous aurons la distribution a priori  $\Pi_i(\theta)$ . Le choix de modéliser le paramètre en question par une loi de  $\beta(a_i, b_i)$ , est expliqué par le fait que celle-ci est fréquemment utilisée si le paramètre à modéliser est borné.

La détermination des paramètres  $a_i$  et  $b_i$  se fait en solvant le système à deux équations non linéaires.

$$S = \begin{cases} F(\lambda_{\min,i}, a_i, b_i) = p \\ F(\lambda_{\max,i}, a_i, b_i) = 1 - p \end{cases} \quad (3.2)$$

NB :L'équation 3.2 est simplement l'inversion d'un intervalle de confiance pour le paramétré  $\lambda$  avec  $p=0.05$ .

Une fois les experts ont formulé leurs avis,nous combinons ces avis afin d'obtenir une distribution a priori unique du paramètre  $\lambda$ . Pour ce faire,nous introduirons deux variables aléatoires comme suit :

- 1  $\Lambda$  : une variable aléatoire continue désignant le taux de défaillance  $\lambda$  de l'équipement.
- 2  $J$  : une variable aléatoire discrète désignant l'expert en fiabilité le plus performant. La fonction de masse de la variable  $J$  est donnée par

$$\mathbb{P}(J = j) = \frac{1}{N} \quad \forall i \in \{1, \dots, N\} \quad (3.3)$$

Nous avons défini la probabilité pour chacun des experts interrogés égal  $\frac{1}{N}$ . Ce choix est justifié par le fait que nous n'avons pas d'information sur la pertinence des avis des experts. On attribue une probabilité égale, afin de ne défavoriser aucun expert. D'où l'utiliser de la fonction de vraisemblance,celui-ci va juger la cohérence des avis des experts. Donc la distribution a priori sur  $\lambda$  pour chaque expert  $J$  sera exprimée par la densité conditionnelle de  $\Lambda$ , étant donné que  $J = j$ . Donc, la distribution a priori est donnée, pour tout  $\lambda \in [0, 1]$  et  $i \in \{1, \dots, N\}$  par

$$\Pi_{\Lambda|J} = \frac{\Gamma(a_j + b_j)}{\Gamma(a_j)\Gamma(b_j)} \lambda^{a_j-1} (1-\lambda)^{b_j-1}, \lambda \in [0, 1] \quad \forall i \in \{1, \dots, N\}, \quad (3.4)$$

$$\begin{aligned} \Pi_{\Lambda}(\lambda) &= \sum_{j=1}^N P(J = j) \Pi_{\Lambda|j} \\ &= \frac{1}{N} \sum_{j=1}^N \frac{\Gamma(a_j + b_j)}{\Gamma(a_j)\Gamma(b_j)} \lambda^{a_j-1} (1-\lambda)^{b_j-1}. \end{aligned} \quad (3.5)$$

Après, nous construisons la distribution a posteriori du paramètre  $\Theta = (\Lambda, J)$ . Afin d'estimer  $\theta = (\lambda, j)$ , nous utiliserons l'inférence bayésienne qui déterminera cette distribution conjointe, combiner avec les données enregistrées sur l'équipement durant

son fonctionnement. Cette distribution est la distribution a posteriori, elle sera ainsi évaluée pour un  $\lambda$  donnée et pour un expert  $i$  donné. Sachant que  $n$  durées de bon fonctionnement de l'équipement, notées par  $x_i$  pour  $i = 1, \dots, n$ , le théorème de Bayes permet d'écrire la distribution a posteriori du paramètre  $\theta$  comme suit

$$\Pi_{\Theta|X}(\theta|x) = \frac{\Pi_{X|\Theta}(x|\theta) \Pi_\Theta(\theta)}{\Pi_X(x)} = \frac{\Pi_{X|\Theta}(x|\theta) \Pi_\Theta(\theta)}{\int_0^\infty \Pi(x|\theta) \Pi(\theta) d\theta}.$$

Le dénominateur de l'équation précédent est une constante de normalisation est indépendante de  $\theta$ . La distribution a posteriori sera proportionnelle au numérateur de cette équation.  $\Pi_{X|\Theta}(x|\theta)$  est la vraisemblance des données  $x$  étant donné le paramètre  $\theta$  et  $\Pi_\Theta(\theta)$  est la distribution a priori sur le paramètre  $\lambda$ . L'équation suivante donne l'expression de la distribution a posteriori sur le paramètre  $\theta$ .

$$\begin{aligned} \Pi_{\Theta|X}(\theta|x) &\propto \Pi_{X|\Theta}(x|\theta) \Pi_\Theta(\theta) \\ &\propto \lambda^n \exp\left(-\lambda \sum_{i=1}^n x_i\right) \Pi_{\Lambda,J}(\lambda, j) \\ &\propto \lambda^n \exp\left(-\lambda \sum_{i=1}^n x_i\right) \mathbb{P}(J = j) \Pi_{\Lambda|J}(\lambda|j). \end{aligned} \tag{3.6}$$

Dans le cas où la distribution a priori et la fonction de vraisemblance sont conjuguées, la forme analytique de la distribution posteriori ne sera pas connue, d'où l'utilisation des méthodes de simulation MCMC afin de calculer cette distribution a postériori puisque la distribution a priori suit la loi bêta et la vraisemblance et exponentielle ne sont pas conjuguées. Pour cela, nous utiliserons le logiciel de statistique R afin de résoudre le problème tel expliqué dans le chapitre 2.

### 3.1.3 Calcul de la loi a priori des avis des experts

Dans notre cas, la forme de la distribution a posteriori sur le paramètre  $\theta$  est inconnue. Dans cette section, nous allons expliquer comment on va utiliser l'algorithme de simulation MCMC. Plus précisément, l'algorithme Metropolis-Hastings pour notre modèle 3.4. La procédure est la suivante

1. Initialiser les valeurs  $\lambda$  et  $j$ .
2. Pour un  $l = 1, \dots, M$ 
  - Générer  $u \sim U[0, 1]$ .
  - Générer  $\theta^* \sim q(\theta^* | \theta^l)$
  - Si le ratio  $u < \alpha$ , alors  $\theta^{l+1} = \theta^*$  sinon  $\theta^{l+1} = \theta^l$ ,

où  $M$  est le nombre d'échantillons et  $U[0, 1]$  est la loi uniforme en 0 et 1.  $\alpha$  est le ratio d'acceptation exprimé par :

$$r = \min \left( 1, \frac{\Pi_{\Theta|X}(\theta^* | x) q(\theta^l | \theta^*)}{\Pi_{\Theta|X}(\theta^l | x) q(\theta^* | \theta^l)} \right),$$

$q(\theta^* | \theta^l)$  est la loi instrumentale. Cette dernière génère un nouveau candidat de  $\theta^*$  à partir de l'ancien valeur de  $\theta^l$  de l'algorithme, on doit générer un  $\lambda$  initial et un expert, dans notre cas la loi utilisé est  $N(\lambda^l, \sigma^2)$ .

Le choix de loi normale est justifié par le fait que cette loi est symétrique. Par conséquent le calcul du ratio  $r$  sera simple à calculer, en plus du choix du paramètre  $\sigma$  de telle façon que l'algorithme converge rapidement vers la distribution a posteriori du paramètre  $\lambda$ . Après le choix de la loi instrumentale, on calcule le ratio  $r$ . En utilisant :

$$q(\theta^* | \theta^l) = q(\theta^l | \theta^*),$$

Ceci nous permettra de simplifier le ratio  $r$  comme suit :

$$r = \min \left( 1, \frac{\Pi_{\Theta|X}(\theta^*|x)}{\Pi_{\Theta|X}(\theta^l|x)} \right) = \min \left( 1, \frac{(\lambda^*)^n \exp(-\lambda^* \sum_{i=1}^n x_i) \mathbb{P}(J = j^*) \Pi_{\Lambda|J}(\lambda^*|j^*)}{(\lambda^l)^n \exp(-\lambda^l \sum_{i=1}^n x_i) \mathbb{P}(J = j^l) \Pi_{\Lambda|J}(\lambda^l|j^l)} \right).$$

Cependant, quand l'algorithme termine ces étapes nous récupérons un vecteur des valeurs du paramètre  $\lambda$ . Nous pouvons ainsi calculer la moyenne et l'écart-type par

$$\bar{\lambda} = \frac{1}{M} \sum_{i=1}^M \lambda^i,$$

et

$$\sigma^2 = \frac{1}{M-1} \sum_{i=1}^M (\lambda^i - \bar{\lambda})^2.$$

Nous avons développé un programme avec le logiciel R qui permet de combiner les avis des experts. Ce programme permet de calculer la fonction a priori des avis des experts, ainsi le choix du bon paramètre de la loi a posteriori. Le code se trouve en Annexe A.

## 3.2 Calcul du DIC et la p-valeur bayésien

Le DIC permet de classer plusieurs modèles. Plus le DIC est petit, plus la qualité d'ajustement du modèle est meilleure et, par conséquent, la distribution a priori sur le paramètre  $\lambda$  est plus vraisemblable. Le DIC se calcule à l'aide de l'échantillon de  $\lambda$  que nous avons obtenu en appliquant l'algorithme de Metropolis-Hastings, les équations suivantes nous résument le calcul de cette quantité :

$$\begin{cases} DIC = 2\bar{D} - D(\bar{\lambda}_{post}) \\ \bar{D} = \frac{1}{M} \sum_{i=1}^M -2\log((\lambda^i)^M e^{(-\lambda^i \sum_{i=1}^n x_i)}) \\ D(\bar{\lambda}_{post}) = -2\log((\bar{\lambda}_{post})^n e^{(-\bar{\lambda}_{post} \sum_{i=1}^n x_i)}) \end{cases}$$

donc le DIC égale a l'équation :

$$DIC = -4 \frac{n}{M} \sum_{l=1}^M \log(\lambda^l) + 2n \log(\lambda_{post}) + 2\lambda_{post} \sum_{i=1}^n x_i, \quad (3.7)$$

La deuxième quantité statistique est le p-value bayésienne. Le calcul de cette quantité utilise l'échantillon obtenu de l'algorithme de Metropolis-Hasting. Critère de sélection du modèle plus la valeur-p est proche de 0.5 le modèle est bon, les étapes du calcule se résume dans les étapes suivantes :

1. Pour un  $l=1$  a  $M$ .
  - Simule un échantillon  $x^{rep,l}$  de taille  $n$  de la loi exponentielle de paramètre  $\lambda$
  - Calcule des deux quantités statistiques  $Dev_{obs}(x, \lambda^l)$  et  $Dev_{rep}(x^{rep,l}, \lambda^l)$ .

$$\begin{cases} Dev_{obs}(x, \lambda^l) = -2\log((\lambda^l)e^{(-\lambda^l x_i)}) \\ Dev_{rep}(x^{rep,l}, \lambda^l) = -2\log((\lambda^l)e^{(-\lambda^l x_i^{rep,l})}) \end{cases}$$

2. Nous calculons la valeur-p Bayésien à l'aide de l'équation

$$p-value = \frac{1}{m} \sum_{l=1}^M \mathbb{1}_{\{Dev_{rep}(x^{rep,l}, \lambda^l) \geq Dev_{obs}(x, \lambda^l)\}} \quad (3.8)$$

Dans le chapitre 4 nous allons présenter les résultats de tous ces quantités pour les modèles utilisés afin de bien voir la différence entre ces méthodes

# Chapitre 4

## ANALYSE ET INTERPRÉTATION DES RÉSULTATS

Dans ce chapitre nous analyserons les résultats obtenus avec les méthodes citées dans le chapitre 2 et 3. En premier lieu nous générerons les observations de bon fonctionnement d'une machine. Pour le paramètre  $\lambda=0.4$ , le nombre d'observations générées est 41. Nous avons supposer que nous avons deux experts qui donnent leurs avis à propos du taux de défaillance  $\lambda$ . Ces deux avis sont considérés indépendants du paramètre  $\lambda$ . L'expert 1 est qualifié, l'expert 2 est moins qualifié que le premier.

## 4.1 Comparaison des modèles de combinaison d'avis des experts

Après la génération des observations de bon fonctionnement nous estimerons le paramètre de cette loi, afin de l'implémenter dans notre algorithme MCMC comme valeur initiale, afin d'estimer la loi a priori des avis des experts, ainsi les paramètres de la loi apriori de chaque expert. Les tableaux suivants présentent les valeurs simulées de bon fonctionnement notées par  $x_i$  et  $\hat{\lambda} = 41/(\sum_1^{41} x_i) = 0.41$ , ainsi que les paramètres de loi de chaque expert.

2.590	5.159	5.694	1.435	0.280	1.994	1.091
0.861	1.782	0.886	3.790	0.296	0.935	0.989
0.805	3.659	2.183	0.838	4.688	3.566	11.711
5.043	4.682	1.237	0.091	0.316	0.092	4.481
1.280	3.396	1.451	1.257	0.816	2.990	2.700
0.602	0.373	0.843	1.779	2.970	0.455	

TABLE 4.1 – Durées de bon fonctionnement.

Expert( $j$ )	$\lambda_{j,min}$	$\lambda_{j,max}$	$a_j$	$b_j$
1	0.3	0.5	25.45350	38.47207
2	0.6	0.76	61.83467	28.79216

TABLE 4.2 – Avis des deux experts et paramètres des distributions a priori.

Les paramètres  $a_j$  et  $b_j$  pour chaque expert  $j$  sont obtenus à partir de la résolution de l'équation 3.2 à l'aide du logiciel R. Par la suite, nous allons étudier trois modèles :

- 1 Modèle 1 : considère l'expert 1 comme la seule source d'information afin d'estimer la fonction a priori.
- 2 Modèle 2 : prend la même pondération pour les deux experts afin de formuler la fonction a priori telle que définie dans l'équation (3.4) (agrégation des avis des experts).
- 3 Modèle 3 : on utilisera pour les paramètres de la loi a priori la moyenne des paramètres pour les deux experts (Méthode de la moyenne).

L'étude des trois modèles nous permettra de quantifier la qualité de la distribution a priori de chacun de ces modèles. Pour choisir le modèle qu'on choisira, nous utiliserons deux quantités statistiques qui vont mesurer la qualité d'ajustement de chaque modèle. Il s'agit du DIC (Deviance Information Criterion) et la valeur-p Bayésien, Section 2.6.1. Le calcul de chaque quantité se fait à l'aide de l'échantillon de  $\lambda$  que nous avons obtenu en appliquant l'algorithme de Metropolis-Hastings présenté dans section 2.4.4.

Le DIC permet de classer plusieurs modèles. En effet, plus le DIC est petit, plus la qualité d'ajustement du modèle est meilleure et par conséquent la distribution a priori sur le paramètre  $\lambda^{simulé}$  est plus proche du paramètre observé. Nous avons procédé au calcul des statistiques des distributions a posteriori pour chaque modèle. Les résultats sont présentés dans le tableau suivant :

Modèle	$\lambda_{MCMC}$	$\sigma_\lambda$	DIC	valeur-p bayésien
1	0.4067	0.0438	155.808	0.58
2	0.403	0.048	156.0369	0.6
3	0.508	0.0457	157.4655	0.18

TABLE 4.3 – Résultats a posteriori des trois modèles.

Selon le tableau 4.3, le modèle d'agrégation bayésienne est le meilleur puisqu'il possède une valeur de DIC la plus petite. Ce modèle peut savoir que l'expert 1 est plus performant que l'expert 2.

Expert	contribution
1	0.92
2	0.08

TABLE 4.4 – Probabilité a priori pour chaque expert du modèle d'agrégation des avis des experts.

En effet, il accorde une probabilité a posteriori plus élevée à l'expert 1 qui est jugé très expérimenté comme le montre le Table 4.4.

D'après la figure 4.1, on constate que le vecteur simulé du paramètre  $\lambda$  est dans l'intervalle  $[0.3, 0.7]$ ; donc il donne un poids important à l'expert 1. La valeur du paramètre de bon fonctionnement est de  $0.4 \in [0.3, 0.7]$ . Par contre la méthode de mélange nous donne une estimation du paramétrisé  $\lambda = 0.508$ , et aussi la valeur 0.4 n'appartient à la plage des valeurs générées par Metropolis-Hasting comme nous le montre la figure 4.2.

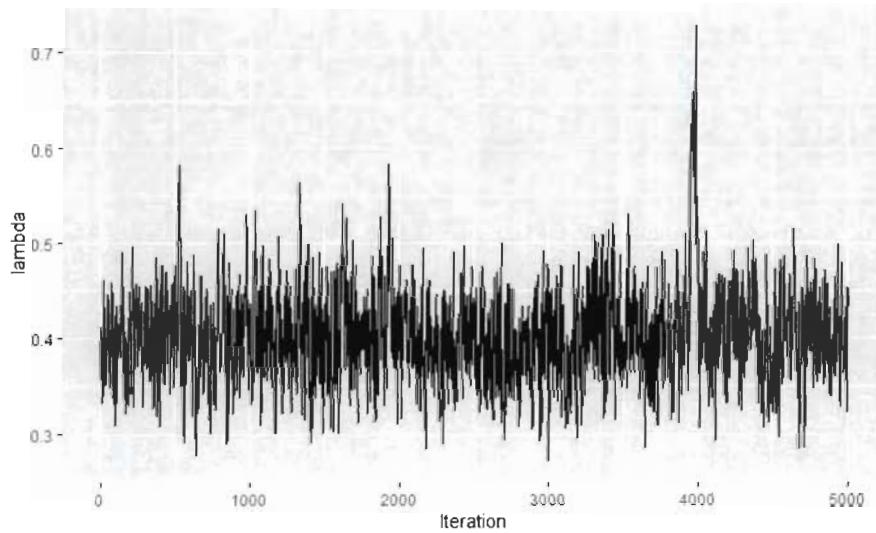


FIGURE 4.1 – Valeurs simulées du paramètre d'agrégation d'avis des experts

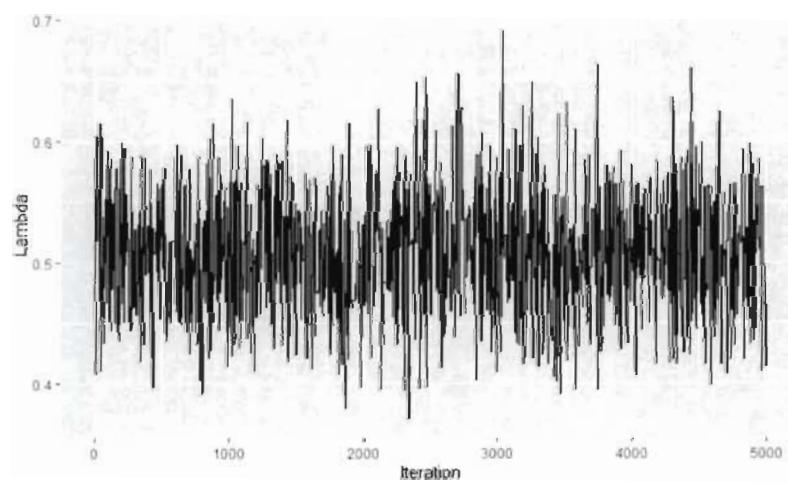


FIGURE 4.2 – Valeurs simulées du paramètre par la méthode agrégation de la moyenne

## 4.2 Validation des résultats d'agrégation des avis des experts

Dans la section précédente, nous avons présenté les résultats des modèles listées dans le chapitre 2. Cette section confirmera notre conclusion à propos du modèle d'agrégation des avis des experts. Nous allons comparer le vecteur simulé à partir de la méthode Metropolis Hasting puis les quantile générer a partir  $n = 5000$  observations et d'une loi bêta avec les paramètres  $a = 25.45350$  et  $b = 38.47207$ .

### 4.2.1 Validation par les quantiles simulés et théoriques

Dans cette sous-section, nous allons présenter le graphique des quantiles observés d'algorithme Metropolis Hasting vs les quantiles générés d'une loi bête(5000,25.45350,38.47207). Comme nous le montre la figure 4.3, nous pouvons conclure graphiquement que les

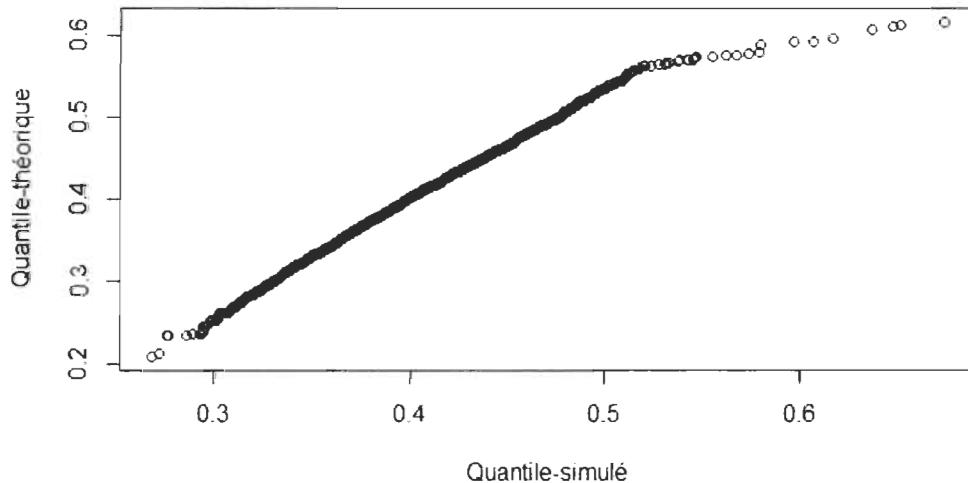


FIGURE 4.3 - Quantile-simulé vs Quantile-théorique

valeur simulées proviennent d'une loi bêta de l'expert 1 jugé expérimente. Par la

suite, nous effectuerons une régression sans ordonné à l'origine afin de montrer que les quantiles simulés sont égaux aux quantiles générés à partir d'une loi bêta avec les paramètres de l'expert 1 jugé qualifié, les résultats de la régression sont résumé dans la figure 4.4.

Le modèle de notre régression explique 99,98 % les quantiles théoriques à partir des quantiles simulés, en plus l'équation de la régression peut s'écrire :

$$\text{Quantile - théorique} = \text{Quantile - simulé} + \epsilon, \quad \epsilon \sim N(0, 1).$$

D'après la figure 4.4, on a le coefficient de la variable Quantile-simulé égale à 1 donc on peut conclure que les quantiles-théoriques et les Quantiles-simulés proviennent de la même loi. Le modèle d'agrégation des avis des expert nous donne une bonne estimation de la fonction a priori, on combinant les observations et la loi de chaque expert.

```
Residuals:
    Min      1Q   Median      3Q      Max
-0.060093 -0.012058 -0.000408  0.009667  0.042025

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
Quantsim 1.0021570  0.0005847   1714 <2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.01668 on 4999 degrees of freedom
Multiple R-squared:  0.9983,    Adjusted R-squared:  0.9983
F-statistic: 2.937e+06 on 1 and 4999 DF,  p-value: < 2.2e-16
```

FIGURE 4.4 – résultat de la régression des quantiles théoriques vs quantile simulé

#### 4.2.2 Validation par les densités des valeurs simulées du paramètre

Dans cette section nous allons présenter les densités simulées par les trois méthodes utilisées. D'après la figure 4.5, on constate que la densité des valeurs générées par

la méthode d'agrégation des avis des experts est similaire, à la densité des valeurs générées par l'expert 1. Par contre, la densité de la méthode moyenne nous donne une densité totalement différente de l'expert 1 et de l'agrégation des avis des experts.

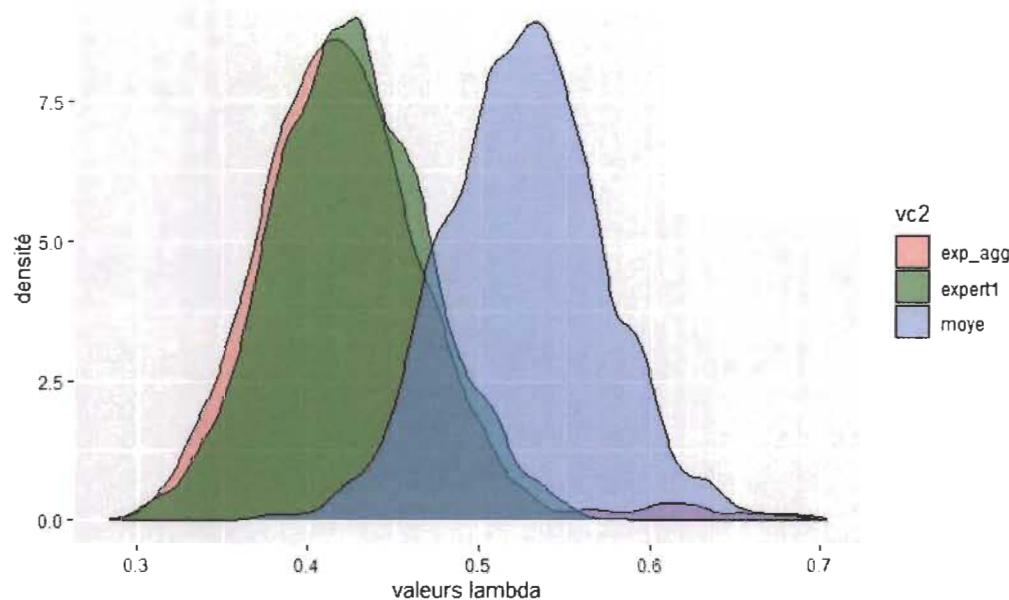


FIGURE 4.5 – Densités des valeurs du paramètre générées par les trois méthodes

### Conclusion :

Pour conclure cette section, nous constatons que le modèle d'agrégation des avis des experts, donne une bonne estimation du paramètre  $\lambda$  en combinant les données observées de bon fonctionnement d'une machine ainsi la fonction a priori sur le paramètre des deux experts. Selon les résultats, cette méthode permet de déterminer l'expert compétent et de lui attribuer un poids très important.

## 4.3 Application du modèle d'agrégation sur les données HQ

### 4.3.1 Illustration des données

Dans ce chapitre nous appliquerons, l'algorithme utilisé dans la section précédente afin, de l'exécuter sur les données fournies par l'équipe du projet PRIAD, du bon fonctionnement d'une machine du réseau transport chez HQ. L'équipe nous a donné trois avis des experts ainsi que le bon fonctionnement observé à partir de 2014 jusqu'à 2018. Ces données sont résumées dans les tableaux suivants :

Expert	MTBF (années)	Taux (annuel)
1	38,46153846	0,026
2	83,33333333	0,012
3	26,31578947	0,038

TABLE 4.5 – Avis des experts sur le bon fonctionnement fourni par l'IREQ.

Année	Taux
2014	0,017
2015	0,023
2016	0,019
2017	0,025
2018	0,018

TABLE 4.6 – Taux de bon fonctionnement observé fourni par IREQ.

### 4.3.2 Interprétation des résultats

Après l'application de l'algorithme nous observons que même si les avis fournis par les experts, ne coïncident pas avec les taux observés. On remarque que le taux estimé par la méthode d'agrégation nous donne des résultats proches des avis des experts ainsi que le taux observé comme nous le montre le tableau 4.7. Donc notre modèle

est capable de trouver un fonction qui combine les informations fourni par plusieurs sources, est une estimation qui permet de tenir en compte les avis des experts ainsi, que les données observées comme nous le montre les figures 4.6, 4.7, 4.8, 4.9 et 4.10.

Années	Taux(agrégation)	intervalle du taux estimé
2014	0.01456	[0.010,0.023]
2015	0.024	[0.0152,0.0328]
2016	0.024	[0.0166,0.0314]
2017	0.038	[0.026,0.047]
2018	0.024	[0.0182,0.029]

TABLE 4.7 – Estimation du taux de fiabilité par la méthode d'agrégation.

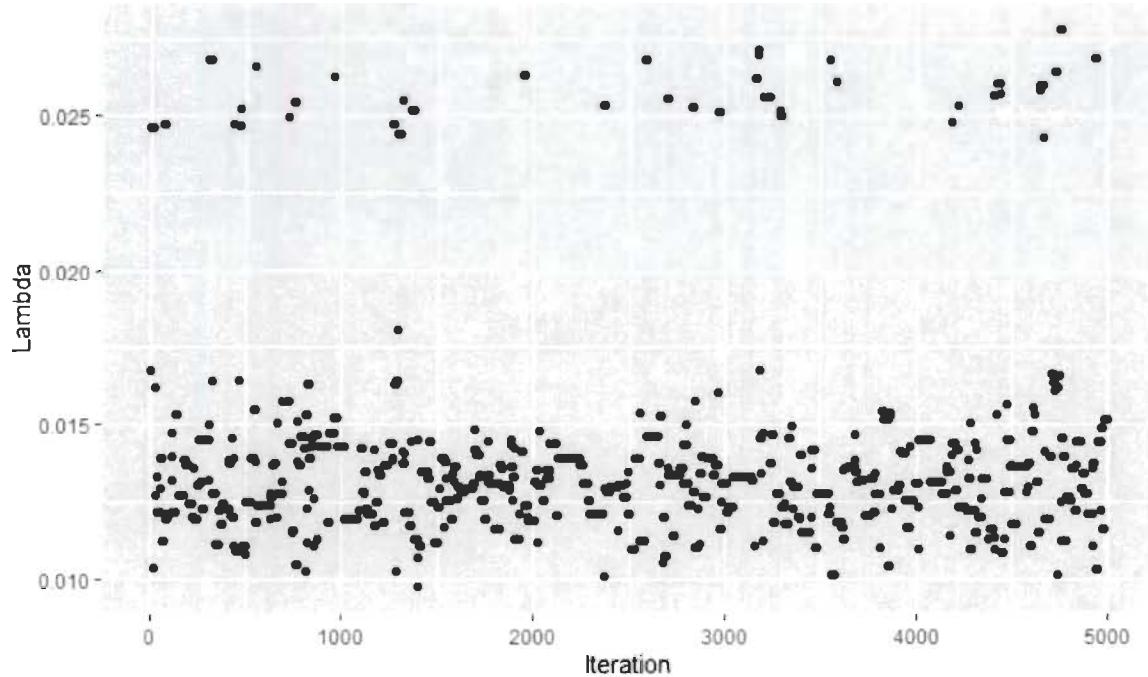


FIGURE 4.6 – Valeurs estimées du taux de fiabilité année 2014 par la méthode D'agrégation

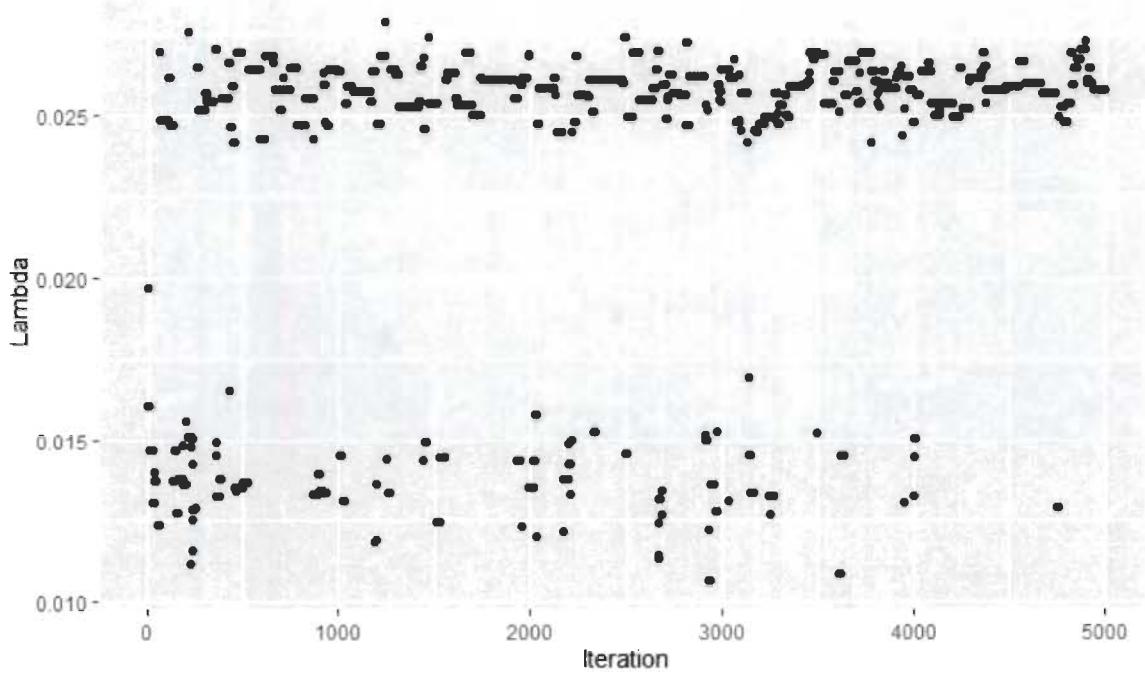


FIGURE 4.7 – Valeurs estimées du taux de fiabilité année 2015 par la méthode D'agrégation

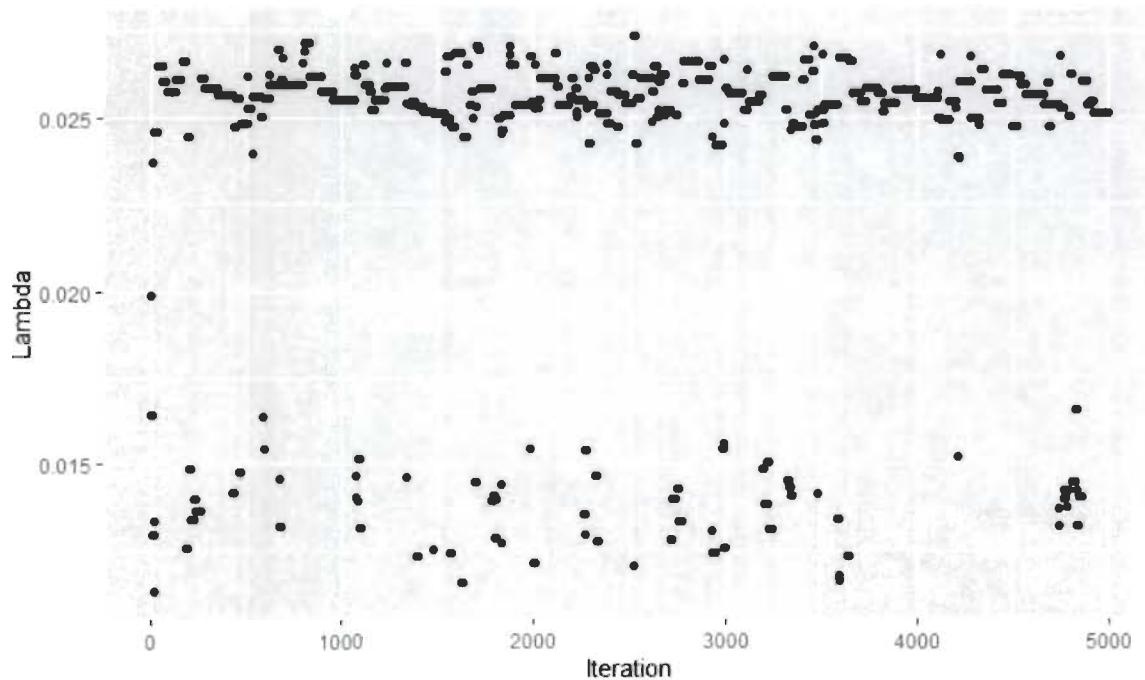


FIGURE 4.8 – Valeurs estimées du taux de fiabilité année 2016 par la méthode D'agrégation

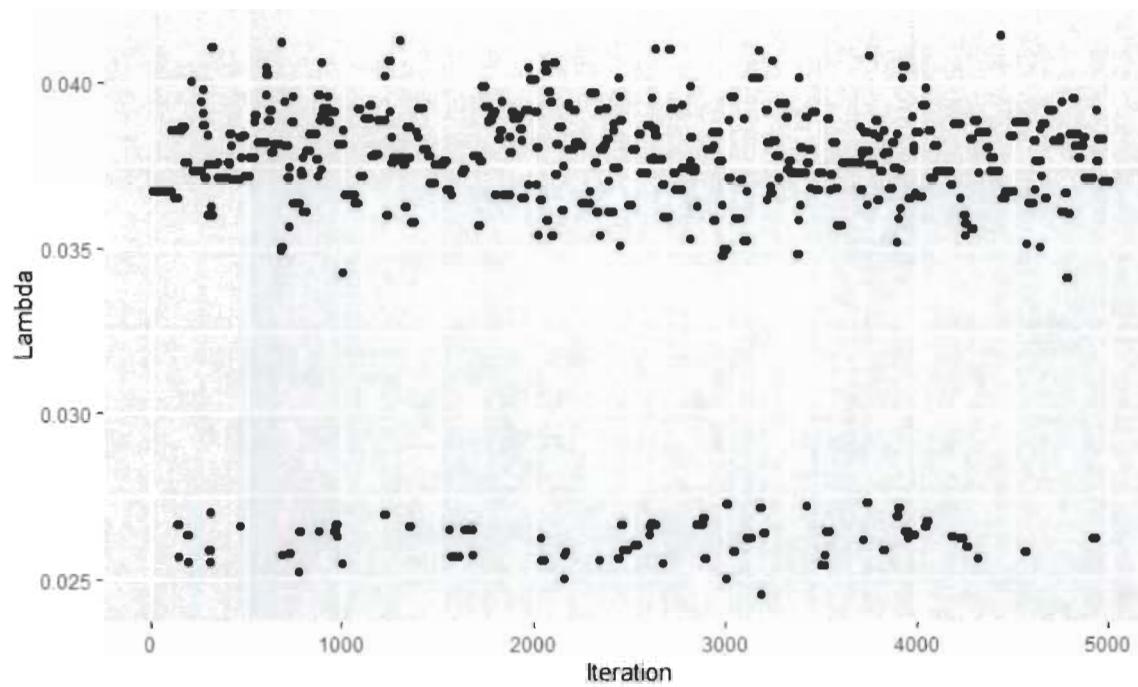


FIGURE 4.9 – Valeurs estimées du taux de fiabilité année 2017 par la méthode D'agrégation

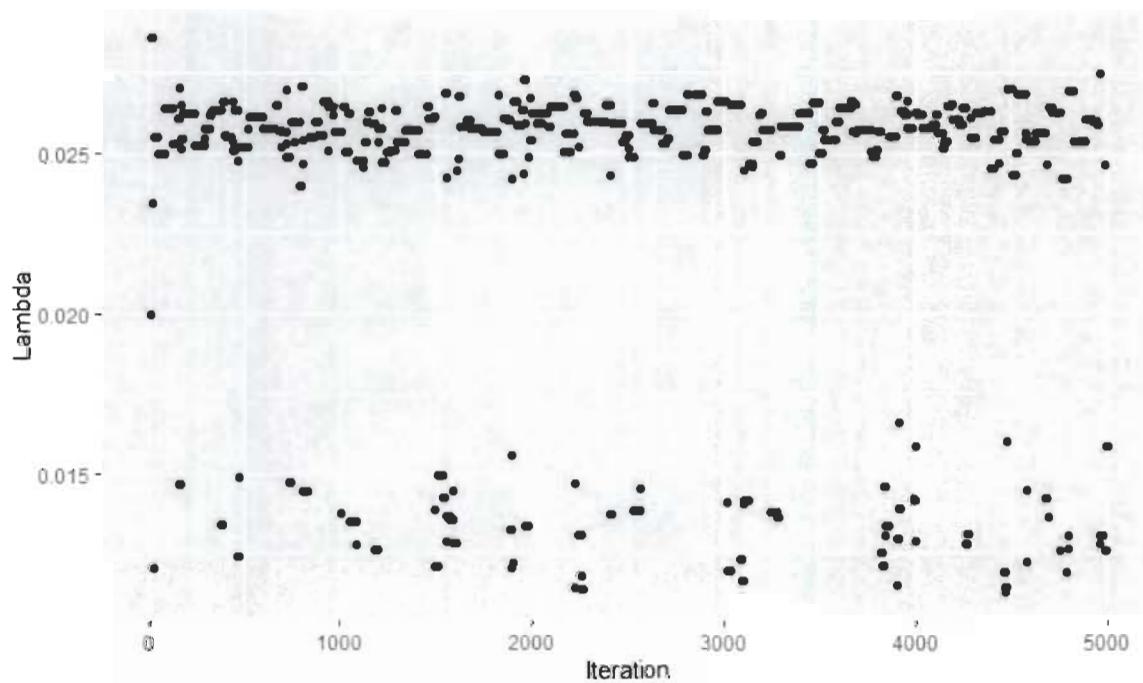


FIGURE 4.10 – Valeurs estimées du taux de fiabilité année 2018 par la méthode D'agrégation

# Chapitre 5

## Conclusion

Ce travail suggère un modèle qui permet de prendre en compte plusieurs avis d'experts provenant d'un même domaine d'expertise, ou de domaines différents. L'inférence Bayésienne permet de combiner des données de défaillance (vraisemblance) avec les avis d'experts, exprimés sous forme de distributions a priori sur un ou plusieurs paramètres inconnus du modèle de défaillance, afin de déduire les distributions a posteriori sur les mêmes paramètres. Contrairement à la méthode du système PRIAD, qui suggère deux estimateurs de fiabilité, puis déterminer le quelle est le plus cohérent aux avis d'experts les plus compétents. Afin de bien modéliser et résoudre notre modèle, nous avons utilisé la méthode de simulation Monte-Carlo par chaînes de Markov, car les distributions a posteriori sur les paramètres d'intérêt ne sont pas connues. L'algorithme utilisé est Metropolis-Hastings. Cependant l'algorithme cité permet la modélisation et la résolution du modèle proposé avec deux variables aléatoires simultanément : le taux de défaillance de l'équipement et la variable désignant l'expert le plus performant. Notre modèle est validé par deux critères :

- 1 Le critère DIC (Deviance information criterion)
- 2 La valeur-P Bayésienne.

Ce modèle et les critères de validation, sont programmés sous le logiciel R. les

données ont été simulées afin de tester l’efficacité de nos méthodes utilisées. En résumé ce travail nous a permis de traiter les points suivants :

- Appliquer les méthodes statistiques bayésiennes, afin d’évaluer le paramètre de fiabilité et des avis des experts en fiabilité.
- Utiliser la méthode de simulation MCMC pour modéliser et estimer des paramètres à partir des lois inconnu.
- Élaborer un programme en R, qui permet de combiner les avis des experts sur les taux de défaillance.

Pour conclure, le modèle d’agrégation des avis des experts présente un avantage majeur, est l’estimation du poids de chaque expert, celant leurs compétences. Même si n’on a pas une idée sur la compétence des experts, il permet aussi d’actualiser l’information décrivant la performance des experts une fois la vraisemblance est connue.

# Bibliographie

- [1] Alain CÔTÉ, Olivier BLANCKE, Stéphane ALARIE, Amira DEMS, Dragan KOMLJENOVIC et Darragi MESSAOUDI : Combining historical data and domain expert knowledge using optimization to model electrical equipment reliability. pages 1–6, 2020.
- [2] Sabin LESSARD : *processus stochastique*. ellipses, 2014.
- [3] Nicolas GRENON-GODBOUT : Méthodes mcmc : amélioration d'un algorithme d'adaptation régionale et applications à la climatologie. -, -:1–39, 2015.
- [4] Edward L KAPLAN et Paul MEIER : Nonparametric estimation from incomplete observations. *Journal of the American statistical association*, 53(282):457–481, 1958.
- [5] G Ronald HERD : Estimation of reliability from incomplete data. In *Proceedings of the 6th national symposium on reliability and quality control*, pages 202–217, 1960.
- [6] Elsayed A ELSAYED : Reliability engineering. *Wiley StatsRef : Statistics Reference Online*, 2012.
- [7] Andrew KS JARDINE et Albert HC TSANG : *Maintenance, replacement, and reliability : theory and applications*. CRC press, 2013.
- [8] Mostafa BACHA, Gilles CELEUX, Edwige IDÉE, André LANNOY et Dominique VASSEUR : *Estimation de modèles de durées de vie fortement censurées*. Éd. Eyrolles, 1998.

- [9] Olivier BLANCKE, Antoine TAHAN, Dragan KOMLJENOVIC, Normand AMYOT, Mélanie LÉVESQUE et Claude HUDON : A holistic multi-failure mode prognosis approach for complex equipment. *Reliability Engineering & System Safety*, 180: 136–151, 2018.
- [10] Nicholas METROPOLIS, Arianna W ROSENBLUTH, Marshall N ROSENBLUTH, Augusta H TELLER et Edward TELLER : Equation of state calculations by fast computing machines. *The journal of chemical physics*, 21(6):1087-1092, 1953.
- [11] Christian P ROBERT, George CASELLA et George CASELLA : *Introducing monte carlo methods with r*, volume 18. Springer, 2010.
- [12] W Keith HASTINGS : Monte carlo sampling methods using markov chains and their applications. 1970.
- [13] Michael A WALLACH, Nathan KOGAN et Daryl J BEM : Group influence on individual risk taking. *The Journal of Abnormal and Social Psychology*, 65(2):75, 1962.
- [14] Anthony O'HAGAN, Caitlin E BUCK, Alireza DANESHKHAH, J Richard EISER, Paul H GARTHWAITE, David J JENKINSON, Jereiny E OAKLEY et Tim RAKOW : *Uncertain judgements : eliciting experts' probabilities*. John Wiley & Sons, 2006.
- [15] Christian GENEST, James V ZIDEK *et al.* : Combining probability distributions : A critique and an annotated bibliography. *Statistical Science*, 1(1):114–135, 1986.
- [16] B. RUMPFF F. et TWARDY L. WINTLE, L. FIDLER : Expert status and performance. *PLoS One*, 6:e22998, 2011.
- [17] David J SPIEGELHALTER, Nicola G BEST, Bradley P CARLIN et Angelika VAN DER LINDE : Bayesian measures of model complexity and fit. *Journal of the royal statistical society : Series b (statistical methodology)*, 64(4):583–639, 2002.
- [18] Sandip SINHARAY et Hal S STERN : Posterior predictive model checking in hierarchical models. *Journal of Statistical Planning and Inference*, 111(1-2):209–221, 2003.

- [19] Isabelle ALBERT, Sophie DONNET, Chantal GUIHENNEUC-JOUYAUX, Samantha Low-CHOY, Kerrie MENGERSSEN, Judith ROUSSEAU *et al.* : Combining expert opinions in prior elicitation. *Bayesian Analysis*, 7(3):503–532, 2012.
- [20] Marco Laumanns. Jiri OSENASK : Bayesien optimization algorithms. *ScienceDirect*, 5:301–320, 2010.

# Annexe A

## Code R :

```
#####Résolution de l'équation{3.1}#####
library(nleqslv)

fun <- function(x) {
  f <- numeric(length(x))
  f[1] <- pbeta(0.6,x[1],x[2])-0.05
  f[2] <- pbeta(0.76,x[1],x[2])-0.95
  f
}
x.start <- c(1,1)
nleqslv(x.start,fun)

###parametre loi apriori d'expert 1 et 2#####
a[1]=25.45350
a[2]=61.83467
b[1]=38.47207
```

```

b[2]=28.79216

a=numeric(2)

b=numeric(2)

#simulation des observation de bon fonctionnement#####
wm=rexp(41,0.4)

lamdam=1/mean(wm) ##Estimateur de maximum de vraisemblance#####

Tom=sum(wm)

#####Estimation du paramétré par MCMC de l'expert 1#####
x=numeric(5000)

x[1]=lamda

for (t in 2:5000) {

y=x[t-1] + 0.2*rnorm(1)

if(runif(1)<(((y/x[t-1])^41)*exp((x[t-1]-y)*To1)*
(dbeta(y,a,b))/(dbeta(x[t-1],a,b))){

x[t]=y

}

else{

x[t]=x[t-1]

}

}

#####Calcule du DIC d'expert 1#####
Moy=mean(x)

dicexp=-4*41*mean(log(x))+2*41*log(Moy)+2*Moy*Tom

#####Valeur-P#####
DEobs=numeric(5000)

DErep=numeric(5000)

for(i in 1:5000){DEobs[i]=-(2*41*log(x[i]))+(2*41*x[i]*mean(w1))

trep=rexp(41,x[i])

DErep[i]=-(2*41*log(x[i]))+(2*41*x[i]*mean(trep))}
```

```

length(which(DErep>DEobs))

##### Agrégation des avis expert#####
#####Estimation du paramétré par MCMC#####

Nsim=5000

X=numeric(5000)

X[1]=lamda # initialisation de la chaine

for (t in 2:Nsim){

y=X[t-1]+0.05*rnorm(1)

rho=((y/X[t-1])^40)*exp((X[t-1]-y)*To1))*  

(0.5*(dbeta(y,a[1],b[1])+dbeta(y,a[2],b[2]))/(0.5*(dbeta(X[t-1],a[1],b[1])+
dbeta(X[t-1],a[2],b[2]))))

X[t]=X[t-1] + (y-X[t-1])*runif(1)<rho)

}

#####DIC agrégation #####
moy=mean(X)

dicexp=-4*41*mean(log(X))+2*41*log(moy)+2*moy*Tom

#####Valeur-P#####
DEobs=numeric(5000)

DErep=numeric(5000)

for(i in 1:5000){DEobs[i]=-(2*41*log(x[i]))+(2*41*x[i]*mean(w1))
trep=rexp(41,x[i])
DErep[i]=-(2*41*log(x[i]))+(2*41*X[i]*mean(trep))}

length(which(DErep>DEobs))

##### méthode Moyenne#####
#####Estimation du paramétré par MCMC#####

Ma=mean(a)
Mb=mean(b)

Z=numeric(5000)

Z[1]=lamda

for (t in 2:5000) {

```

```

y=Z[t-1]+ 0.2* rnorm(1)

if (runif(1)<(((y/Z[t-1])^41)*exp((Z[t-1]-y)*To1))*  

(dbeta(y,Ma,Mb)/(dbeta(Z[t-1],Ma,Mb)))){  

Z[t]=y  

}  

else{  

Z[t]=Z[t-1]  

}  

}  

#####DIC Moyenne#####
Moy=mean(Z)  

dicepx=-4*41*mean(log(Z))+2*41*log(Moy)+2*Moy*To1  

##### Valeur-P#####
DEobs=numeric(5000)  

DErep=numeric(5000)  

for(i in 1:5000){DEobs[i]=-(2*41*log(Z[i]))+(2*41*Z[i]*mean(w1))  

trep=rexp(41,x[i])  

DErep[i]=-(2*41*log(x[i]))+(2*41*Z[i]*mean(trep))}  

length(which(DErep>DEobs))
#####

```