

République Algérienne Démocratique et Populaire
Ministère de l'Enseignement Supérieur et de la Recherche
Scientifique

Université de A/Mira de Bejaia
Faculté des Sciences Exactes
Département mathématiques



Mémoire présenté en vue de l'obtention du diplôme de Master
En mathématiques
Option : Statistique et Analyse Décisionnelle

THEME

Modélisation de l'information a priori
et applications en fiabilité

Présenté par :

HAMOUCHE Tinhinane & SEBBANE Souhila

Soutenu, le 17/06/2015 devant le jury composé de :

Mr BOURAINE. M	Président	MAA	U.A/MIRA, Bejaia
M ^{me} LAGHA. K	Rapporteur	MCB	U.A/MIRA, Bejaia
M ^{me} HOCINE. S	Examinatrice	MAB	U.A/MIRA, Bejaia
M ^{elle} BRAHEM. H	Examinatrice	Doctorante	U.A/MIRA, Bejaia

*******Dédicaces*******

Voilà on est arrivé à la fin de ce travail que je dédie à toutes les personnes qui comptent pour moi.

A mes très chers parents à qui je présente ma profonde gratitude pour tout ce qu'ils ont fait pour moi que DIEU les protège.

A mon cher frère unique Ali que je l'estime, il est toujours à mes côtés avec son aide et son soutien.

Et à mes très chères sœurs qui m'ont encouragé durant tout au long de mon travail de recherche.

Et à toute ma famille Hamouche.

A mes très chères copines de résidence berchiche 1 et ma cousine Amel.

A mon très cher binôme que je respecte Souhila, que j'estime beaucoup, à qui je souhaite que la réussite et le bonheur.

Ainsi toute la promotion SAD.

Tout particulièrement mon futur époux Nassim qui est toujours devant moi durant ces trois dernières années.

HAMOUCHE Tinhinane

******Dédicaces******

Je dédie ce modeste travail:

*A mes chers parents : ma mère et mon père qui m'ont
encouragé tout au long de mes études.*

*A mes deux frères : Abdesslam et messaoud, et ma sœur
nesserine a qui je souhaite la réussite dans leurs études..*

A toute ma famille.

A ma meilleure amie et ma chère binôme Tinhinane.

*Je dédie également ce mémoire à tous mes amis.
A toute la promotion SAD.*

SEBBANE Souhila

N° d'ordre : 28/2004-M/MT

République Algérienne Démocratique et Populaire
**MINISTÈRE DE L'ENSEIGNEMENT SUPÉRIEUR ET DE LA
RECHERCHE SCIENTIFIQUE**
UNIVERSITÉ DE BEJAIA A. MIRA
FACULTÉ DES SCIENCES EXACTES
Département Mathématique
Mémoire présenté pour l'obtention du diplôme de Master
en mathématiques

Spécialité : Statistique et Analyse Décisionnelle



THÈME

Modélisation de l'information a priori et applications en fiabilité

Présenté par :

HAMOUCHE Tinhinane et SEBBANE Souhila

Soutenu , le 16/06/2015 devant le jury composé de :

Mr BOURAINE. M	Président :	MCB	U.A/MIRA Bejaia
M ^{me} LAGHA. K	Rapporteur :	MCB	U.A/MIRA Bejaia
M ^{elle} BRAHEM. H	Examinatrice :		U.A/MIRA Bejaia
M ^{me} Hocine. S	Examinatrice :	MAA	U.A/MIRA Bejaia

Remerciements

Il nous a été très difficile d'écrire cette page par souci d'oublier les nombreuses personnes qu'il faut citer pour leur aide, accueil et soutien. Qu'elles soient toutes assurées de notre plus profonde reconnaissance même si leurs noms n'y figurent pas.

Éloge à Allah, le tout puissant et miséricordieux qui nous a prodigué la santé, le courage et la patience afin de mener ce travail à son terme.

Ce travail a été dirigé par Docteur Mme LAGHA.K à qui nous adressons tout particulièrement nos vives reconnaissances et gratitudes, nous la remercions pour son soutien, son aide, sa disponibilité permanente, ses encouragements constants, ses conseils judicieux, et surtout ses critiques rigoureuses qui nous ont été très précieuses pour structurer ce travail et nous pousser à réaliser ce qui n'a pas encore été accompli.

Nous adressons nos vifs remerciements à Mr Bouraine.M pour l'honneur qu'il nous a fait en acceptant de présider le jury et d'évaluer ce mémoire à sa juste valeur.

Nous remercions M^{elle} Braham.H pour avoir accepté de juger et examiner ce travail, ainsi que son aide durant notre travail de recherche.

Nos vifs remerciements vont également à Mme Hocine.S pour avoir accepté d'examiner notre travail et de l'enrichir avec ses propositions.

Nous tenons à remercier chaleureusement notre encadreur de stage Mr GUECHARI.E, qui nous a aidé tout au long de notre stage pratique, avec sa pédagogie, sa disponibilité et sa prévenance, ce qui nous a facilité le travail, et tous les membres d'équipes du service planification et ordonnancement ainsi que ceux du service méthode.

Table des matières

Liste des tableaux	2
Liste des figures	3
Liste des abréviations	4
Introduction	5
1 Généralités sur la Statistique Bayésienne et la Fiabilité	7
1.1 Généralités sur la Statistique Bayésienne	7
1.1.1 Principe de l'Approche Bayésienne	8
1.1.2 Information a priori	9
1.1.3 Loi a priori	9
1.1.4 Modèle Bayésien	9
1.1.5 Loi des observations	10
1.1.6 Distribution a posteriori	10
1.1.7 Le mécanisme de l'inférence bayésienne	13
1.1.8 Loi a priori impropre	13
1.1.9 Les bases de la théorie de décision	14
1.1.10 Minimaxité et Admissibilité	16
1.1.11 Estimateur de Bayes	18
1.1.12 Test et intervalle de crédibilité	20
1.2 Généralités sur la fiabilité	22
1.2.1 Concepts élémentaires de la fiabilité	23
1.2.2 Les indices de fiabilité	26
1.2.3 Matériel réparable	31

1.2.4	Matériel irréparable	31
1.2.5	Fiabilité des systèmes	32
1.2.6	Principaux modèles de fiabilité	35
2	Modélisation et choix de la loi a priori	46
2.1	Lois a prioris non informatives	47
2.1.1	Loi a priori de Laplace	47
2.1.2	Loi a priori de Jeffreys	50
2.1.3	Loi a priori invariante	53
2.1.4	Loi a priori de référence	57
2.1.5	Loi a priori coïncidante - (matching prior)	58
2.2	Lois a prioris informatives	59
2.2.1	Loi a priori conjuguée	59
2.2.2	Loi a priori d'entropie maximale	62
2.2.3	Approche hiérarchique	64
2.2.4	Loi subjective	65
2.3	Le poids de l'a priori dans la réponse bayésienne	66
3	Estimation Bayésienne des paramètres de la fiabilité	70
3.1	Les sources d'informations en fiabilité	71
3.2	Description du système	72
3.3	Modélisation du nombre de défaillances	73
3.4	Modélisation exponentielle des temps de défaillances	80
3.5	Modélisation Weibull des temps de défaillances	88
3.6	Application aux données du stage	96
3.6.1	Collecte et constitution d'échantillons de données	96
3.6.2	Traitement statistique des données	96
3.6.3	Estimation des paramètres de la fiabilité	97
	Conclusion Générale	100
	Annexes	102
	Bibliographie	112

Liste des tableaux

2.1	Loi a priori conjuguée naturelle pour quelques familles exponentielles usuelles..	61
2.2	Exemples de lois à maximum d'entropie.....	63
2.3	Information a priori sur le paramètre p_i pour différentes machines.....	66
2.4	Modèle a priori sur le paramètre p_i correspondant à l'information a priori.....	66
3.1	Données du REX sur les groupes électrogènes.....	77
3.2	Les résultats générés pour les estimateurs fréquentistes et bayésiens.....	79
3.3	Comparaison des intervalles de confiance du taux de défaillance.....	88
3.4	Comparaison des intervalles de crédibilité de θ	95
3.5	Comparaison des intervalles de crédibilité de λ	100

Liste des figures

1.1	Principe de la démarche bayésienne.....	09
1.2	Les différents types de maintenances selon la norme EN 13306 (2001).....	25
1.3	Allure de la fonction de répartition $F(t)$	27
1.4	Allure de la fonction de fiabilité $R(t)$	27
1.5	Allure de l'évolution du taux de défaillance « courbe en baignoire».....	29
1.6	Durées moyennes associées à un matériel.....	30
1.7	Chronogramme d'un équipement réparable.....	31
1.8	Chronogramme d'un équipement irréparable.....	31
1.9	Diagramme de fiabilité d'un système en série.....	32
1.10	Diagramme de fiabilité d'un système en parallèle.....	34
1.11	Diagramme de fiabilité pour un système k parmi n	35
1.12	Densité de probabilité de la loi exponentielle en fonction de λ	36
1.13	Propriété sans mémoire de la loi exponentielle.....	37
1.14	Le taux de défaillance d'une Weibull pour différentes valeurs de β	39
1.15	Densité de probabilité de la loi Weibull.....	39
1.16	Densité de probabilité de la loi Gamma pour différentes valeurs de α et β	42
1.17	La loi Binomiale (n, p)	44
3.1	Distributions a priori et a posteriori de la probabilité de défaillance.....	79
3.2	Ecart sur l'évaluation de la probabilité de défaillance entre l'inférence classique et bayésienne.....	80
3.3	Ajustement des temps de défaillances de la machine souffeuse par la loi exponentielle.....	97

Liste des abréviations

ddl	degrés de liberté
EIREDA	European Industry Reliability Data
Ent	Entropie
IC	Intervalle de confiance
REX	Retour d'expérience
MV	Maximum de vraisemblance
v.a	Variable aléatoire

Introduction

Le fondement de la théorie bayésienne s'est basé sur le théorème d'inversion des probabilités connu sous le nom du théorème de Bayes (1763). Laplace (1774) approfondit la notion de probabilité inverse par l'introduction des probabilités des causes des événements.

La méthode bayésienne consiste à baser l'estimation des paramètres sur une distribution nommée loi a posteriori qui mélange l'information portée par les données et une information a priori exogène aux données (avis d'expert, informations régionales, etc).

Le formalisme bayésien possède plusieurs avantages par rapport à une analyse classique. Tout d'abord, l'introduction de connaissances a priori est susceptible d'améliorer l'estimation des paramètres. De plus, il permet une analyse complète des incertitudes: échantillonnage (c'est le cas également de l'analyse classique), mais aussi de la modélisation (comment choisir la loi a priori ?).

La statistique bayésienne connaît une croissance exponentielle avec des applications dans tous les domaines: science de l'ingénieur (fiabilité), science expérimentale, science sociale, médecine, économie,...etc. Elle est devenue un outil incontournable pour la modélisation des différents phénomènes.

Au début du $XX^{\text{ème}}$ siècle, des enjeux économiques liés au fort développement industriel ont conduit les ingénieurs à s'approprier à des méthodes bayésiennes et les appliquées en fiabilité afin d'améliorer leurs pratiques de façon plus systématique. L'essor de l'informatique a finalement donné lieu à la mise en place de nombreuses techniques permettant de trouver des solutions numériques aux équations représentant les lois de la nature. La combinaison des différents facteurs permet aux concepteurs d'avoir une expérience sur les matériels dont ils connaissent a priori le comportement.

Les ingénieurs utilisent alors, de façon croissante, des méthodes statistiques pour prévoir le comportement de leurs matériels sous l'angle fiabiliste. L'appel aux seules données de REX ou d'essais ne suffit pas pour prévoir le comportement des matériels, il est alors nécessaire de faire appel à une connaissance a priori. Cette connaissance est en général apportée par des experts industriels, et cela permet à l'ingénieur d'estimer de façon satisfaisante les grandeurs typiques de la fiabilité.

Plusieurs travaux de recherche sur la statistique bayésienne en fiabilité ont été réalisés [Procaccia et al. 1992a], [Lannoy and Procaccia, 1994a], [Henri Procaccia et Marie-Christine Sunher, 2003], [Nicolas BOUSQUET, 2006], etc.

L'objectif de ce travail est de répondre aux questions suivantes:

- Comment traduire une information a priori en une loi a priori ?
- Comment traduire la qualité de cette information ?
- Comment choisir la loi a priori quand on dispose et quand on ne dispose pas d'information ?
- Quel est l'impact de cette information dans l'estimation ?

Ce mémoire est constitué de trois chapitres:

Le premier chapitre est un rappel de quelques notions préliminaires de la statistique Bayésienne et de la fiabilité.

Le deuxième chapitre présente les différentes approches de construction d'une loi a priori dans le cas de la disponibilité et la non disponibilité de l'information a priori.

Le troisième est constitué de trois applications en fiabilité. Nous avons modélisé le nombre de défaillance dans le cadre d'un retour d'expérience binomiale où on a estimé la probabilité de défaillance à la sollicitation, par la suite nous avons considéré le cas d'un matériel dans sa période de jeunesse, exprimé par le modèle exponentiel, ensuite le cas d'un modèle weibull modélisant les différents stades de vie de ce matériel et nous avons estimé les paramètres de fiabilité par la méthode bayésienne.

La troisième application concerne la modélisation des durées de vie d'une machine (souffleuse) et l'estimation de ses paramètres de fiabilité avec les données du stage effectué à IFRI (Bejaïa).

Généralités sur la Statistique Bayésienne et la Fiabilité

1.1 Généralités sur la Statistique Bayésienne

Le troisième millénaire sera, dit-on celui de l'information. Aussi, la statistique y sera-t-elle appelée à jouer un rôle important dans le traitement des données, et le paradigme bayésien plus que tout autre, puisqu'il offre un cadre de raisonnement bien adapté à l'intégration des opinions et des faits de toute provenance qui interviennent dans la gestion des risques et la prise de décision dans un contexte d'incertitude. Etant donné un modèle paramétrique statistique d'observations faiblement informatives, issu d'une variable aléatoire (*v.a.*) X distribuée selon $f(x|\theta)$, dépendant d'un paramètre inconnu θ appartenant à un espace Θ de dimension finie, et x la réalisation (valeur observée) de X produite par l'expérience. À l'ajout d'une autre connaissance du comportement de l'échantillon, si elle existe, on espère alors, à condition de pouvoir modéliser raisonnablement cette information, en tirer une précision suffisante de l'estimation de la fonction d'intérêt de θ . Une façon courante d'agglomérer ces deux connaissances et de faire appel aux techniques bayésiennes.

D'une part si toute l'information fournie par l'observation X est contenue dans l'échantillon, on a recours à l'approche classique en utilisant la méthode du maximum de vraisemblance. D'autre part, s'il s'agit du cas où nous avons des connaissances a priori sur le paramètre inconnu θ , le paramètre est considéré comme une variable aléatoire et nous lui attribuons

une loi de probabilité π sur Θ . On applique dans ce cas des méthodes d'estimations bayésiennes.

On introduit dans ce paragraphe le principe de l'approche bayésienne et les différents aspects théoriques de la statistique bayésienne, ainsi que les bases de la théorie de décision.

1.1.1 Principe de l'Approche Bayésienne

Le principe de l'analyse statistique bayésienne est de modéliser des paramètres inconnus de la distribution d'échantillonnage à travers une structure probabiliste. Elle vise à exploiter le plus efficacement possible l'information apportée par X sur le paramètre θ , pour ensuite construire des procédures d'inférence sur ce paramètre. Bien que x ne soit qu'une réalisation aléatoire d'une loi gouvernée par θ , elle apporte une actualisation aux informations préalablement recueillies par l'expérimentateur. Le principe fondamental de l'approche bayésienne part d'hypothèses a priori [46], servant à définir la probabilité subjective, supposée être un concept personnalisé dépendant à la fois des connaissances disponibles a priori de l'expert (degré d'incertitude) ou d'un retour d'expérience déjà disponible, portés sur la valeur de la quantité inconnue θ , et des informations objectives, (i.e. observations nouvelles recueillies) exprimées par la fonction de vraisemblance conditionnée par l'information a priori, qui va confirmer, infirmer ou réactualiser cette connaissance a priori.

La combinaison de ces deux informations, considérées comme des *v.a* qui bien souvent ne sont pas homogènes, permet d'obtenir une densité de probabilité enrichie intégrant toute la connaissance disponible, et a la particularité de présenter une plus faible incertitude que la distribution de probabilité obtenue par la méthode du maximum de vraisemblance.

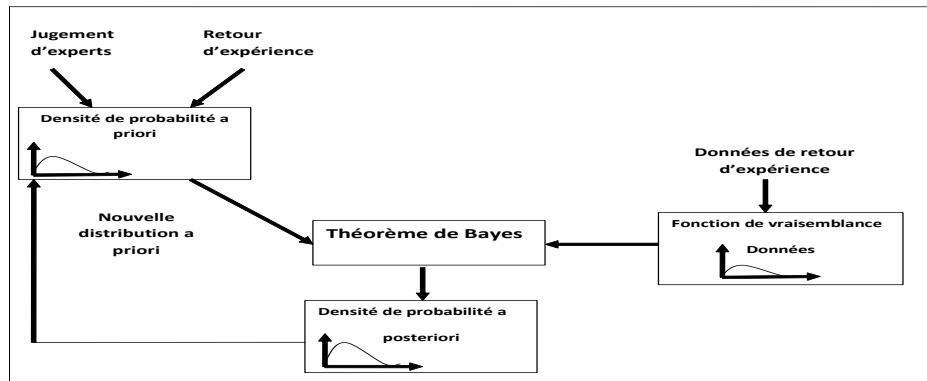


Figure.1.1- Principe de la démarche bayésienne.

1.1.2 Information a priori

On appelle information a priori sur le paramètre θ , toute information disponible sur θ en dehors de celle apportée par les observations. Cette information est entachée d'incertitude (si ce n'était pas le cas, le paramètre θ serait connu avec certitude et on n'aurait pas à l'estimer).

Il est naturel de modéliser cette information à travers une loi de probabilité, appelée loi a priori.

1.1.3 Loi a priori

On considère que l'incertitude sur le paramètre θ d'un modèle peut être décrite par une distribution de probabilité $\pi(\theta)$ sur Θ . L'appellation a priori exprime le fait qu'elle a été établie préalablement à l'observation des données.

1.1.4 Modèle Bayésien

Considérons un échantillon $\underline{X} = (X_1, \dots, X_n)$ de taille n issu d'une variable aléatoire X distribuée selon la densité $f(x|\theta)$ dépendant d'un paramètre inconnu θ . Un modèle statistique bayésien est constitué d'un modèle statistique paramétrique $f(x|\theta)$ et d'une distribution a priori pour les paramètres, notée $\pi(\theta)$.

1.1.5 Loi des observations

L'information fournie par l'observation X est contenue dans la densité $f(x|\theta)$ que l'on représente classiquement, à partir des données de l'échantillon, sous la forme inversée de vraisemblance, interprétée comme loi conditionnelle de X sachant θ .

On note la densité de la loi des observations (la vraisemblance) par $L(\theta, \underline{x})$ définie par:

- Cas discret

$$L(\theta, \underline{x}) = P_\theta(\underline{x}) = P(\underline{X} = \underline{x}|\theta) = \prod_{i=1}^n P(X_i = x_i|\theta). \quad (1.1.1)$$

- Cas continu

$$L(\theta, \underline{x}) = f(\underline{x}|\theta) = \prod_{i=1}^n f(x_i|\theta). \quad (1.1.2)$$

1.1.6 Distribution a posteriori

La distribution a posteriori de θ , notée $\pi(\theta|\underline{x})$, représente la distribution conditionnelle de θ sachant les observations $\underline{x} = (x_1, \dots, x_n)$. Elle donne l'information dont on dispose sur θ , après observation, et elle représente un compromis entre l'information a priori (donnée par la fonction $\pi(\theta)$) et l'information tirée de l'échantillon (donnée par la fonction $L(\theta, \underline{x})$).

Détermination de la loi a posteriori

Théorème de Bayes

Soit $\{(A_i), i = 1, \dots, n\}$ un système complets d'événements définis sur Ω , l'espace des événements. La probabilité de l'événement A conditionnellement à la réalisation de l'événement B , est exprimée par la relation suivante [6]:

$$P(A_i|B) = \frac{P(B|A_i) P(A_i)}{\sum_i P(B|A_i) P(A_i)}. \quad (1.1.3)$$

Cette équation est une conséquence triviale de la définition de la probabilité conditionnelle, et est appelée formule de Bayes.

En utilisant le théorème de Bayes, la distribution a posteriori de θ peut être calculée à partir de la relation suivante:

$$\pi(\theta|\underline{x}) = \frac{\varphi(\theta, \underline{x})}{m(\underline{x})}. \quad (1.1.4)$$

Où, $\varphi(\theta, \underline{x})$ est la loi jointe de θ et \underline{x} , définie par:

$$\varphi(\theta, \underline{x}) = L(\theta, \underline{x}) \pi(\theta), \quad (1.1.5)$$

et $m(x)$ la loi marginale de \underline{x} , avec,

$$m(\underline{x}) = \int_{\Theta} \varphi(\theta, \underline{x}) d\theta = \int_{\Theta} L(\theta, \underline{x}) \pi(\theta) d\theta. \quad (1.1.6)$$

La distribution a posteriori est alors donnée par:

$$\pi(\theta|\underline{x}) = \frac{L(\theta, \underline{x}) \pi(\theta)}{\int_{\Theta} L(\theta, \underline{x}) \pi(\theta) d\theta}. \quad (1.1.7)$$

Cette formule est la mise à jour de la connaissance a priori sur le paramètre à estimer θ , en une connaissance a posteriori $\pi(\theta|\underline{x})$.

Raisonnement proportionnel

Il est parfois très difficile de calculer l'intégrale: $\int_{\Theta} f(\underline{x}|\theta) \pi(\theta) d\theta$. Il est possible d'éviter ce calcul, en utilisant le raisonnement proportionnel.

Définition 1.1.1 Soient deux fonctions réelles f et g définies sur le même espace χ . On dit que f et g sont proportionnelles et on note $f \propto g$, s'il existe une constante a tel que $f(x) = ag(x)$ pour tout $x \in \chi$.

Remarques 1.1.1 1). La relation de proportionnalité est une relation d'équivalence. En particulier, elle est transitive i.e,

$$si (f \propto g) \text{ et } (g \propto h) \text{ alors } (f \propto h), \quad (1.1.8)$$

où f, g, h sont des fonctions

2). Soit f la densité d'une v.a X de loi inconnue, et g la densité d'une v.a Y de loi connue, notée P . Si $f \propto g$ alors $X \sim P$.

3). Dans le contexte bayésien on a [16]:

$$\pi(\theta|\underline{x}) = aL(\theta, \underline{x})\pi(\theta). \quad (1.1.9)$$

$$\text{Avec, } a = \frac{1}{m(\underline{x})} \text{ et } m(\underline{x}) = \int_{\Theta} L(\theta, \underline{x})\pi(\theta) d\theta.$$

D'où,

$$\pi(\theta|\underline{x}) \propto L(\theta, \underline{x})\pi(\theta). \quad (1.1.10)$$

Exemple 1.1.1 Soit X_1, \dots, X_n , un n -échantillon de loi Bernoulli (θ) issu d'une v.a X .

$$L(\underline{x}, \theta) = \prod_{i=1}^n P(X = x_i|\theta) = \theta^{\sum_{i=1}^n x_i} (1 - \theta)^{n - \sum_{i=1}^n x_i}, \theta \in (0, 1).$$

Considérons que θ suit une loi a priori Bêta (a, b), définie par:

$$\pi(\theta) = \frac{1}{B(a, b)} \theta^{a-1} (1 - \theta)^{b-1} I_{[0,1]}(\theta) \text{ avec } a > 0, b > 0,$$

avec $B(a, b)$ est la fonction bêta donnée par:

$$B(a, b) = \frac{\Gamma(a)\Gamma(b)}{\Gamma(a+b)}.$$

La loi a posteriori est alors donnée par proportionnalité,

$$\begin{aligned} \pi(\theta|\underline{x}) &= L(\underline{x}, \theta) \pi(\theta) \\ &\propto (1 - \theta)^{b-1} \theta^{\sum_{i=1}^n x_i} (1 - \theta)^{n - \sum_{i=1}^n x_i} \theta^{a-1} I_{[0,1]}(\theta) \\ &\propto \theta^{\sum_{i=1}^n x_i + a - 1} (1 - \theta)^{n - \sum_{i=1}^n x_i + b - 1} \end{aligned}$$

On reconnaît bien une loi Bêta de paramètres $(a + \sum_{i=1}^n x_i ; b + n - \sum_{i=1}^n x_i)$. C'est-à-dire,

$$\theta|\underline{x} \sim \text{Bêta}(a + \sum_{i=1}^n x_i, b + n - \sum_{i=1}^n x_i).$$

1.1.7 Le mécanisme de l'inférence bayésienne

Le processus d'inférence comporte les étapes suivantes:

- choix d'un modèle relatif aux données (distribution conditionnelle) qui doit appartenir à une famille de lois de probabilité ;
- détermination d'une distribution a priori du paramètre, en intégrant des informations subjectives. Elle est établie à partir des connaissances que l'on a sur le phénomène étudié: essais antérieurs, avis d'experts, retour d'expériences déjà disponible, enquête, etc ;
- obtention de la distribution a posteriori par application du théorème de Bayes ;
- estimation des paramètres du phénomène étudié à partir de la distribution a posteriori. [46].

1.1.8 Loi a priori impropre

Une loi impropre est une application π sur l'espace des paramètres Θ , qui vérifie :

$$\int_{\Theta} \pi(\theta) d\theta = +\infty \quad (1.1.11)$$

Les lois impropres sont obtenues lorsqu'on dispose des critères subjectifs ou théoriques sur la distribution a priori du paramètre, qui conduisent à une mesure de probabilité [20].

Remarques 1.1.2 1) Dans le cas d'une loi impropre, $\pi(\theta)$ n'est pas une densité de probabilité.

2) L'usage des lois a priori impropres est justifié tant que la loi a posteriori est propre (loi de probabilité), car l'inférence bayésienne est basée sur la loi a posteriori.

Exemple 1.1.2 Soit θ un paramètre inconnu de la loi a priori de Haldane, définie par:

$$\pi(\theta) = [\theta(1-\theta)]^{-1} 1_{]0,1[}.$$

$\pi(\theta)$ est une loi a priori impropre. En effet,

$$\begin{aligned} \int_0^1 \pi(\theta) d\theta &= \int_0^1 [\theta(1-\theta)]^{-1} d\theta \\ &= (\ln \theta|_0^1) + (-\ln(1-\theta)|_0^1) = +\infty - (-\infty) = +\infty \end{aligned}$$

1.1.9 Les bases de la théorie de décision

L'objectif général d'une étude statistique est de fournir une décision au statisticien. Le problème le plus dominant auquel on s'intéresse ici est celui d'un individu plongé dans un environnement donné (nature) et qui, sur la base d'observations, est conduit à mener des actions et à prendre des décisions qui auront un coût.

Les espaces intervenant dans l'écriture d'un modèle de décision sont:

- χ : l'espace des observations ou résultats expérimentaux ;
- Θ : l'espace des paramètres ;
- D : l'ensemble des décisions possibles δ .

L'inférence consiste à choisir une règle de décision δ dans D concernant le paramètre $\theta \in \Theta$ sur la base d'une observation $x \in \chi$, avec x et θ sont liés par la loi $f(x|\theta)$.

Règle de décision : Une règle de décision $\delta(x)$ est une fonction, $\delta : x \rightarrow d = \delta(x)$, telle que pour chaque observation x on peut prendre une décision $d = \delta(x)$. Dans le cadre de l'estimation, la règle de décision est un estimateur.

Fonction de coût

La fonction de perte (ou de coût), notée $l(\theta, \delta(x))$, associée à chaque décision δ et à chaque état de la nature θ est une évaluation numérique des pertes provoquées par la décision.

La fonction de coût l est une fonction de $\Theta \times D$ à valeurs dans \mathbb{R}_+ .

Les fonctions de coûts les plus connues sont [43] :

a) Coût quadratique

Introduite par Légende (1805) et Gauss (1810), ce coût est le critère d'évaluation le plus commun. Fondant sa validité sur l'ambiguïté de la notion d'erreur dans un contexte statistique (soit erreur de mesure, soit variation aléatoire), il a aussi donné lieu à de nombreuses critiques. La plus fréquente étant sans doute le fait que le coût quadratique pénalise trop fortement les grandes erreurs. Le coût quadratique est défini par:

$$l(\theta, \delta) = (\theta - \delta)^2. \quad (1.1.12)$$

b) Coût absolu

Une solution alternative au coût quadratique en dimension un est d'utiliser le coût absolu qui est défini par:

$$l(\theta, \delta) = |\theta - \delta|. \quad (1.1.13)$$

c) Coût 0-1

La fonction de coût 0 – 1 est définie par:

$$l(\theta, \delta) = \begin{cases} 1 - \delta & \text{si } \theta \in \Theta_0 \\ \delta & \text{sinon.} \end{cases} \quad (1.1.14)$$

Ce coût est surtout utilisé dans l’approche classique des tests d’hypothèses proposées par Neyman et Pearson. Plus généralement, c’est un exemple typique d’un coût non quantitatif. En effet, pour ce coût, la pénalité associée à un estimateur δ est 0 si la réponse est correcte et 1 sinon.

Fonctions de risque

a) Risque fréquentiste

On appelle fonction de risque fréquentiste, notée $R(\theta, \delta)$, le coût moyen du coût d’une règle de décision, c’est-à-dire [44],

$$R(\theta, \delta) = E[l(\theta, \delta(\underline{x}))] = \int_{\mathcal{X}} l(\theta, \delta(\underline{x})) f(\underline{x}|\theta) dx. \quad (1.1.15)$$

b) Risque a posteriori

On appelle fonction de risque a posteriori notée $\rho(\pi, \delta|\underline{x})$, la moyenne du coût par rapport à la loi a posteriori. C’est-à-dire [26],

$$\rho(\pi, \delta|\underline{x}) = E^\pi[l(\theta, \delta(\underline{x})|\underline{x})] = \int_{\Theta} l(\theta, \delta) \pi(\theta|\underline{x}) d\theta. \quad (1.1.16)$$

c) Risque de Bayes

Soit δ un estimateur de θ et l la fonction de perte associée. On appelle risque bayésien noté $r(\pi, \delta)$, le risque fréquentiste moyenné sur les valeurs de θ selon leur distribution a priori [43]. C’est-à-dire,

$$r(\pi, \delta) = E^\pi[R(\theta, \delta)] = \int_{\Theta} R(\theta, \delta) \pi(\theta) d\theta,$$

D’où,

$$r(\pi, \delta) = \int_{\Theta} \int_{\mathcal{X}} l(\theta, \delta(\underline{x})) f(\underline{x}|\theta) d\underline{x} \pi(\theta) d\theta. \quad (1.1.17)$$

Théorème 1.1.1 *Un estimateur minimisant le risque de Bayes $r(\pi, \delta)$ peut être obtenu en sélectionnant, pour chaque $x \in \mathcal{X}$, la valeur $\delta(\underline{x})$ qui minimise la perte a posteriori $\rho(\pi, \delta|\underline{x})$ [15]. D'où,*

$$r(\pi, \delta) = \int_{\mathcal{X}} \rho(\pi, \delta|\underline{x}) m(\underline{x}) d\underline{x}, \quad (1.1.18)$$

Preuve. En utilisant le Théorème de Fubini nous permettant d'interchanger l'ordre des intégrales, nous obtenons:

$$\begin{aligned} r(\pi, \delta) &= E^{\pi} [R(\theta, \delta)] \\ &= \int_{\Theta} R(\theta, \delta) \pi(\theta) d\theta \\ &= \int_{\Theta} \int_{\mathcal{X}} l(\theta, \delta(\underline{x})) f(\underline{x}|\theta) d\underline{x} \pi(\theta) d\theta \\ &= \int_{\Theta} \int_{\mathcal{X}} l(\theta, \delta(\underline{x})) \pi(\theta|\underline{x}) m(\underline{x}) d\theta d\underline{x} \\ &= \int_{\mathcal{X}} \left[\int_{\Theta} (l(\theta, \delta(\underline{x})) \pi(\theta|\underline{x}) d\theta) \right] m(\underline{x}) d\underline{x} \\ &= \int_{\mathcal{X}} \rho(\pi, \delta|\underline{x}) m(\underline{x}) d\underline{x}. \end{aligned}$$

D'où le résultat. ■

1.1.10 Minimaxité et Admissibilité

Estimateurs randomisés

Définition 1.1.2 *Pour un ensemble de décisions D , on définit D^* comme l'ensemble des probabilités sur D . L'estimateur $\delta^* \in D^*$ est appelé estimateur randomisé.*

Théorème 1.1.2 *(Christian P. Robert [43])*

Pour toute distribution a priori π sur Θ , le risque de Bayes pour l'ensemble des estimateurs randomisés est le même que celui pour l'ensemble des estimateurs non randomisés, soit

$$\inf_{\delta \in D} r(\pi, \delta) = \inf_{\delta^* \in D^*} r(\pi, \delta^*) = r(\pi). \quad (1.1.19)$$

Minimaxité

Le critère de minimaxité que nous présentons maintenant apparaît comme une assurance contre le pire, car il vise à minimiser le coût moyen dans le cas le moins favorable.

Définition 1.1.3 On appelle *risque minimax* associé à la fonction de coût L , la valeur \bar{R} , définie par:

$$\bar{R} = \inf_{\delta \in D} \sup_{\theta} R(\theta, \delta) = \inf_{\delta \in D^*} \sup_{\theta} E_{\theta}[l(\theta, \delta(x))].$$

On appelle *estimateur minimax*, tout estimateur (éventuellement randomisé) δ_0 , tel que

$$\sup_{\theta} R(\theta, \delta_0) = \bar{R}. \quad (1.1.20)$$

Admissibilité

Le critère d'admissibilité induit un ordre partiel sur D^* en comparant les risques fréquentistes $R(\theta, \delta)$ des estimateurs.

Définition 1.1.4 Un estimateur δ_0 est dit *inadmissible* s'il existe un estimateur δ_1 qui le domine. C'est-à-dire, pour tout θ ,

$$R(\theta, \delta_0) \geq R(\theta, \delta_1), \quad (1.1.21)$$

et pour au moins une valeur θ_0 du paramètre,

$$R(\theta_0, \delta_0) > R(\theta_0, \delta_1). \quad (1.1.22)$$

Sinon, δ_0 est dit *admissible*.

Proposition 1.1.1 (Christian P.Robert 2006)

Si δ_0 est admissible de risque constant, δ_0 est l'unique estimateur minimax.

1.1.11 Estimateur de Bayes

Soit $\underline{X} = (X_1, \dots, X_n)$ un n-échantillon de variables aléatoires de densité $f(x|\theta)$, où θ est un paramètre inconnu (à estimer).

Définition 1.1.5 Un estimateur de Bayes associé à une distribution a priori π et une fonction de coût L , est un estimateur δ^π minimisant le risque de Bayes $r(\pi, \delta)$. C'est-à-dire, pour chaque $x \in \mathcal{X}$ [37],

$$\delta^\pi = \arg \min_{\delta \in D} \rho(\pi, \delta | \underline{x}). \quad (1.1.23)$$

Définition 1.1.6 *Estimateur MMSE (Minimum Mean Squar Error)*

On appelle estimateur MMSE tout estimateur bayésien de θ , notée $\hat{\theta}_{MMSE}$, qui minimise la perte quadratique [26].

Proposition 1.1.2 L'estimateur de Bayes δ^π associé à la loi a priori π , et à la perte quadratique est la moyenne a posteriori [43]. C'est-à-dire,

$$\hat{\theta}_{MMSE} = E(\theta | \underline{x}). \quad (1.1.24)$$

Preuve. On a:

$$\begin{aligned} \rho(\pi, \delta | \underline{x}) &= E^\pi [l(\theta, \delta(\underline{x}) | \underline{x})] \\ &= \int_{\Theta} l(\theta, \delta) \pi(\theta | \underline{x}) d\underline{x} \\ &= \int_{\Theta} (\theta - \delta)^2 \pi(\theta | \underline{x}) d\theta \\ &= \int_{\Theta} \theta^2 \pi(\theta | \underline{x}) d\theta - 2\delta \int_{\Theta} \theta \pi(\theta | \underline{x}) d\theta + \delta^2 \int_{\Theta} \pi(\theta | \underline{x}) d\theta. \\ &= E^\pi(\theta^2 | \underline{x}) - 2\delta E^\pi(\theta | \underline{x}) + \delta^2. \end{aligned}$$

La décision δ qui minimise le risque a posteriori $\rho(\pi, \delta | \underline{x})$ est celle qui vérifie :

$$\frac{d}{d\delta} \rho(\pi, \delta | \underline{x}) = 0.$$

Ce qui donne

$$-2E^\pi(\theta | \underline{x}) + 2\delta = 0.$$

Et donc,

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

■

Définition 1.1.7 *Estimateur de MAP (Maximum A Posteriori)*

On appelle estimateur MAP, l'estimateur qui maximise la loi a posteriori $\pi(\theta|\underline{x})$.
C'est-à-dire,

$$\hat{\theta}_{MAP} = \arg \max_{\theta} (\pi(\theta|\underline{x})). \quad (1.1.25)$$

L'estimateur est alors le mode de la loi a posteriori.

Remarques 1.1.3 1) L'estimateur MAP est l'estimateur de Bayes associé au coût $0 - 1$ [15].

2) L'estimateur de Bayes associé au coût L_1 est la médiane a posteriori [43].

3) Les estimateurs de Bayes les plus utilisés en pratique sont le MAP et le MMSE.

4) On retiendra la moyenne de la loi a posteriori comme estimateur de Bayes obtenu, relativement au coût quadratique.

Définition 1.1.8 *Estimateur de Bayes Généralisé*

Soit $\pi(\theta)$ une loi a priori impropre, sous la condition que $\int_{\Theta} L(\underline{x}, \theta) \pi(\theta) d\theta$ est convergente, et $\pi(\theta|\underline{x})$ la loi a posteriori définie par la relation (1.1.5).

On appelle estimateur de Bayes généralisé de θ noté $\hat{\theta}_G$, la moyenne de cette loi a posteriori relativement à la fonction de coût quadratique, qui minimise pour chaque x la perte a posteriori [15].

Exemple 1.1.3 Soit le modèle Bayésien défini par:

$X|\theta \sim \exp(\theta)$ et $\theta \sim \text{Uniforme sur } \mathbb{R}^+$.

On a :

$$X|\theta \sim \exp(\theta) \Rightarrow f(x|\theta) = \theta \exp(-\theta x), x > 0.$$

D'où,

$$L(\theta, \underline{x}) = \prod_{i=1}^n f(x_i|\theta) = \prod_{i=1}^n [\theta \exp(-\theta x_i)] = \theta^n \exp\left(-\theta \sum_{i=1}^n x_i\right).$$

Et, $\theta \sim \text{Uniforme}(\mathbb{R}^+) \Rightarrow \pi(\theta) = 1_{\mathbb{R}^+}(\theta) = \begin{cases} 1, & \text{si } \theta > 0 \\ 0, & \text{sinon,} \end{cases}$ est une loi a priori impropre.

D'où la loi a posteriori est:

$$\begin{aligned} \pi(\theta|\underline{x}) &\propto L(\theta, \underline{x}) \pi(\theta) = \left[\theta^n \exp\left(-\theta \sum_{i=1}^n x_i\right) \right] 1_{]0, +\infty[}(\theta), \theta > 0, \\ &\Rightarrow \theta|\underline{x} \sim \text{Gamma}\left(n+1, \sum_{i=1}^n x_i\right). \end{aligned}$$

L'estimateur de Bayes généralisé associé à la fonction de coût quadratique est alors,

$$\hat{\theta}_G = E(\theta|\underline{x}) = \frac{n+1}{\sum_{i=1}^n X_i}.$$

Propriétés de l'estimateur bayésien

- Si la distribution a priori π est strictement positive sur Θ de risque de Bayes fini, et la fonction de risque $R(\theta, \delta)$ est une fonction continue de θ pour tout δ , l'estimateur de Bayes δ^π est admissible [43].

- L'estimateur de bayes est biaisé et convergent en probabilité.

- La loi a posteriori peut être asymptotiquement approximée par une loi normale $N(E(\theta|\underline{x}), V(\theta|\underline{x}))$.

1.1.12 Test et intervalle de crédibilité

Un test statistique peut être considéré soit comme un procédé statistique, c'est-à-dire une fonction définie sur l'espace des observations à valeurs dans un espace à deux éléments que l'on appelle "accepter" et "rejeter" une hypothèse, ou comme une façon pour le statisticien de gérer ses doutes relatifs à son modèle statistique. Un test bayésien est basé sur la loi a posteriori.

Intervalle de crédibilité

Définition 1.1.9 (*Région α -crédible*).

Pour $0 < \alpha < 1$, une région α -crédible pour θ est un sous-ensemble $C \in \Theta$ tel que

$$P^\pi(\theta \in C | X = x) = 1 - \alpha, \tag{1.1.26}$$

habituellement C est un intervalle. Il existe une infinité de régions α -crédible, il est donc logique de s'intéresser à celle qui a un volume minimal. Pour cela, nous allons introduire la notion d'une région HPD.

Définition 1.1.10 (*Région HPD, Highest Posterior Density*)

La région HPD est la région C_α^π définie par:

$$C_\alpha^\pi = \{\theta \in \Theta, \text{ tel que } \pi(\theta|X = x) \geq h_\alpha\}. \quad (1.1.27)$$

Où,

$$h_\alpha = \sup h \text{ tel que } P^\pi \{(\theta, \pi(\theta|x) \geq h|x)\} \geq 1 - \alpha.$$

Les tests

On suppose que l'espace Θ des paramètres est partitionné en Θ_0 et Θ_1 , et que $P(\theta \in \Theta_k) > 0$, $k = 0, 1$. On souhaite tester:

$$H_0 : \theta \in \Theta_0 \text{ contre } H_1 : \theta \in \Theta_1.$$

En statistique bayésienne, la réponse à un tel test repose sur les probabilités a posteriori des hypothèses H_0 et H_1 , données par:

$$P(H_0|x) = P(\theta \in \Theta_0|x) = \int_{\Theta_0} \pi(\theta|x) d\theta \text{ et } P(H_1|x) = 1 - P(H_0|x). \quad (1.1.28)$$

En pratique, une hypothèse (H_0 ou H_1) est acceptée dès que sa probabilité a posteriori est jugée suffisamment forte (typiquement supérieur à 0.9 ou à 0.95). Il se peut qu'aucune des probabilités a posteriori n'excède 0.9. Deux attitudes sont alors possibles:

- soit on ne prend pas de décision, et on décide par exemple de recueillir davantage d'observations ;
- soit on choisit l'hypothèse dont la probabilité a posteriori est la plus grande.

Considérons les probabilités a priori de chacune des deux hypothèses H_0 et H_1 :

$$P(H_0) = P(\theta \in \Theta_0) = \int_{\Theta_0} \pi(\theta) d\theta \text{ et } P(H_1) = P(\theta \in \Theta_1) = \int_{\Theta_1} \pi(\theta) d\theta. \quad (1.1.29)$$

Le facteur de Bayes

Considérons le test suivant:

$$H_0 : \theta \in \Theta_0 \text{ contre } H_1 : \theta \in \Theta_1.$$

La loi a priori est donnée par le mélange de lois suivantes:

$$\pi(\theta) = P(H_0) \pi_0(\theta) + P(H_1) \pi_1(\theta), \quad (1.1.30)$$

où, $\pi_i(\theta) = P(\theta|H_i)$ est la probabilité a priori de θ sous l'hypothèse H_i et $P(H_1) = 1 - P(H_0)$.

La fonction de vraisemblance sous H_i s'écrit:

$$L(\underline{x}|H_i) = \int_{\theta \in \Theta_i} L(\underline{x}|H_i, \theta) \pi_i(\theta) d\theta, i = 0 \text{ ou } 1.$$

Définition 1.1.11 *Le facteur de Bayes de l'hypothèse H_0 en faveur de l'hypothèse H_1 s'exprime par:*

$$B_F = \left[\frac{P(H_0)}{P(H_1)} \right] \left[\frac{P(H_0|\underline{x})}{P(H_1|\underline{x})} \right]. \quad (1.1.31)$$

Le facteur de Bayes est un indicateur objectif de l'hypothèse $H_0 : \theta \in \Theta_0$.

Interprétation du facteur de Bayes

Comme la règle de Bayes consiste à choisir l'hypothèse ou le modèle de plus grande probabilité a posteriori, le facteur de Bayes peut être interprété comme suit:

Pour

- $B_F > 1$, nous favorisons H_0 ,
- $10^{-\frac{1}{2}} \leq B_F < 1$, la certitude de H_0 est fautive est minimale,
- $10^{-1} \leq B_F \leq 10^{-\frac{1}{2}}$, cette certitude est substantielle,
- $10^{-2} \leq B_F \leq 10^{-1}$, la certitude est forte,
- $B_F \leq 10^{-2}$, la certitude est décisive et nous devons rejeter H_0 .

On constate ici, un autre point de différence entre l'approche classique et l'approche bayésienne et un test bayésien ne consiste pas à accepter une hypothèse et rejeter l'autre, mais bien à comparer les deux hypothèses et prendre la meilleure.

1.2 Généralités sur la fiabilité

L'évaluation du niveau de fiabilité des équipements nécessite une parfaite connaissance du comportement opérationnel de ces derniers. La modélisation de la fiabilité des équipements s'obtient à partir de l'analyse des données du retour d'expérience du matériel en exploitation, ou par des essais (tests) sur un échantillon de composants ou du matériel lui même.

1.2.1 Concepts élémentaires de la fiabilité

Systeme

Un système (system, en anglais) est un ensemble déterminé d'éléments discrets (ou composants) interconnectés ou en interactions.

On propose la classification suivante [48]:

(pièce) \subset (composant) \subset (système élémentaire) \subset (sous système) \subset (système) .

Défaillance

La défaillance (failure, en anglais) correspond à l'altération ou la cessation de l'aptitude d'un bien à accomplir une fonction requise.

La défaillance peut être intrinsèque ou primaire, inhérente aux faiblesses du dispositif ou extrinsèque par le fait que son origine est dû à une cause extérieure (cas d'une mauvaise utilisation du matériel).

Différents critères peuvent être pris en considération pour distinguer les défaillances [30]:

- mode de défaillance ;
- cause de la défaillance ;
- fréquence de la défaillance ;
- rapidité de la manifestation ;
- degré de gravité ;
- conséquences ou effets de la défaillance.

Durée de vie

La durée de vie est une variable aléatoire notée par T qui représente le temps écoulé depuis la mise en service du dispositif à l'instant $t = 0$ jusqu'à l'instant de sa première défaillance [29].

Fiabilité

La fiabilité (Reliability, en anglais) est l'aptitude d'un système à maintenir les fonctionnalités pour lesquelles il a été conçu, pendant une durée de vie déterminée et dans des conditions données. Un système fiable est un système qui ne tombe pas en panne pendant sa durée de mission [46].

La notion de fiabilité se distingue selon l'étape étudiée de la vie du système (ou de l'équipement):

- La fiabilité prévisionnelle

C'est la fiabilité déterminée durant la phase de conception sur la base d'un modèle mathématique défini à partir des données de fiabilité de ses composants.

- La fiabilité estimée

C'est la fiabilité mesurée après avoir conçu le système et ceci à l'aide d'un ensemble d'essais.

- La fiabilité opérationnelle

C'est la fiabilité mesurée sur un matériel en exploitation. Elle dépend des conditions réelles d'exploitation et du support logistique.

Maintenance

La maintenance (maintenance, en anglais) désigne l'ensemble des opérations nécessaires pour maintenir un système à un niveau de fiabilité donné. Sa maintenabilité caractérise la facilité avec laquelle on le maintient en bon état de fonctionnement.

Il existe deux grands types de maintenance: la maintenance corrective et la maintenance préventive [20].

Les types de maintenance

- **Maintenance préventive :** Elle est réalisée pendant l'état de marche du matériel.

Les interventions ont pour but de réduire la dégradation du matériel et de retarder une éventuelle panne future. On distingue parmi les interventions de maintenance préventive:

- maintenance préventive systématique ;
- maintenance préventive planifiée ;

- maintenance préventive conditionnelle.

- **Maintenance corrective** : Elle est réalisée à la suite d'une défaillance, et qui a pour but de réaliser le fonctionnement du matériel. On procèdera à une réparation, ou à un remplacement du matériel défaillant, suivant que l'on ait affaire à des matériels complexes réparables, ou au remplacement global de l'entité élémentaire (composant) défaillante. On distingue parmi les interventions de maintenance corrective :

- maintenance curative ;
- maintenance palliative.

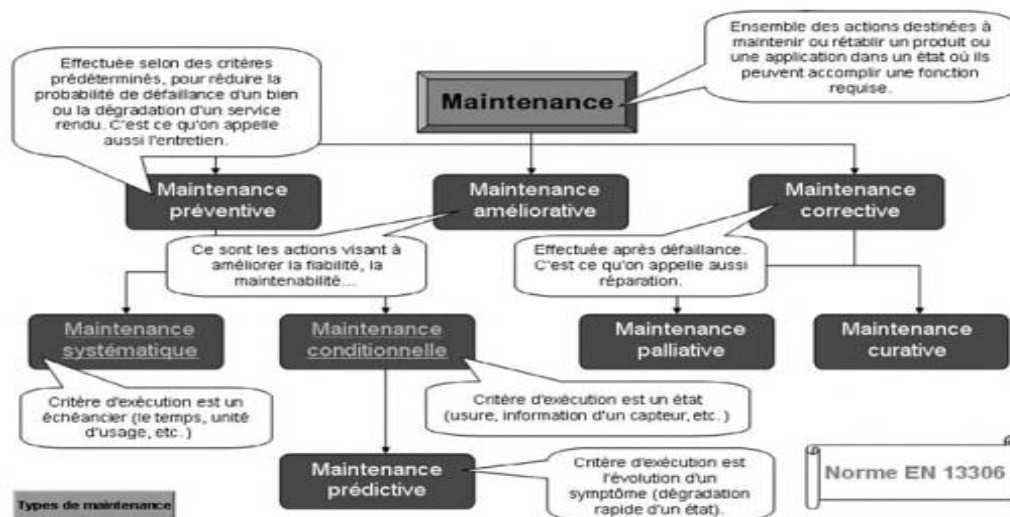


Figure.1.2- Les différents types de maintenances selon la norme EN 13306 (2001)[19].

Maintenabilité

La maintenabilité (Maintainability en anglais) est l'aptitude d'une entité à être maintenue ou rétablie dans un état dans lequel elle peut accomplir une fonction requise, lorsque la maintenance est réalisée dans des conditions données avec des procédures et des moyens prescrits. Elle est caractérisée par la probabilité $M(t)$ que l'entité E soit en état, à l'instant t , d'accomplir ses fonctions, sachant que l'entité était en panne à l'instant 0. C'est-à-dire [46],

$$M(t) = P(\text{l'entité est réparée sur } [0, t]). \quad (1.2.1)$$

Disponibilité

La disponibilité (Availability en anglais) est l'aptitude d'une entité à être en état d'accomplir les fonctions requises dans des conditions données et à un instant donné. Elle est caractérisée par la probabilité $A(t)$ que l'entité E soit, à l'instant t , en état d'accomplir les fonctions requises dans des conditions données. C'est-à-dire [46],

$$A(t) = P(\text{l'entité est non défaillante à l'instant } t). \quad (1.2.2)$$

1.2.2 Les indices de fiabilité

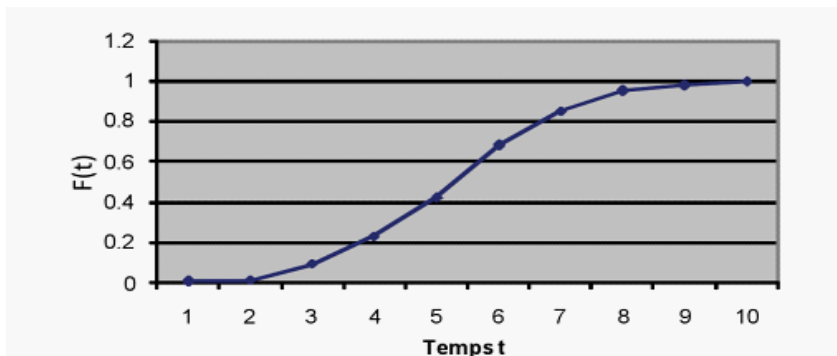
La fonction de répartition

La fonction de répartition, notée $F(t)$, désigne la fonction de probabilité de panne (ou de défaillance) d'un composant ou d'une structure. Elle représente la probabilité d'avoir au moins une défaillance avant le temps t . Elle est exprimée par [22]:

$$F(t) = P(T < t) = \int_0^t f(u) du, \quad t > 0, \quad (1.2.3)$$

où f est la densité de probabilité de défaillance à l'instant t . On a [30]:

$$\begin{aligned} f(t) &= \frac{dF(t)}{dt} = \lim_{dt \rightarrow 0} \frac{F(t+dt) - F(t)}{dt}, \quad t > 0 \\ &= \lim_{dt \rightarrow 0} \frac{P(t < T \leq t + dt)}{dt} \\ &\Rightarrow f(t) dt \simeq P(t < T \leq t + dt). \end{aligned} \quad (1.2.4)$$

Figure.1.3- Allure de la fonction de répartition $F(t)$.

Fonction de fiabilité

Elle est généralement notée $R(t)$ et représente la probabilité de non défaillance dans l'intervalle de temps de $[0, t]$, c'est-à-dire la probabilité de défaillance au-delà du temps t . C'est la fonction complémentaire de la fonction de répartition.

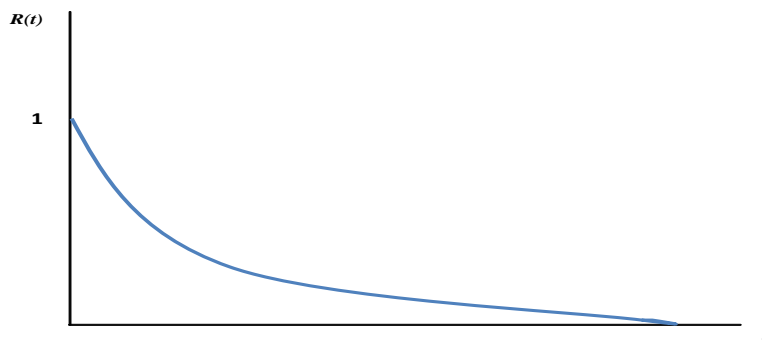
$$R(t) = P(T \geq t) = 1 - F(t) = 1 - \int_0^t f(u) du, \quad t > 0. \quad (1.2.5)$$

Propriété

La fonction de fiabilité $R(t)$ est une fonction monotone et décroissante sur $[0 ; t]$.

Elle vérifie [21]:

$$R(0) = 1, \quad 0 \leq R(t) \leq 1 \quad \text{et} \quad \lim_{t \rightarrow \infty} R(t) = 0.$$

Figure.1.4- Allure de la fonction de fiabilité $R(t)$.

Taux de défaillance

Le taux de défaillance $\lambda(t)$ désigne la proportion des dispositifs qui, ayant survécus jusqu'à l'instant t , ne sont plus en vie à l'instant $t + \Delta t$, parmi les dispositifs en vie à l'instant t . Autrement dit, c'est la probabilité conditionnelle pour qu'un élément tombe en panne au cours de $[t; t + \Delta t]$, sachant qu'il a fonctionné sans défaillance sur $[0, t]$ [27].

Le taux de défaillance s'exprime par:

$$\begin{aligned}\lambda(t) &= \lim_{\Delta t \rightarrow 0} \frac{P((t < T \leq t + \Delta t) / T > t)}{\Delta t} \\ &= \lim_{\Delta t \rightarrow 0} \frac{1}{\Delta t} \left[\frac{P(t < T \leq t + \Delta t)}{P(T > t)} \right] \\ &= \frac{1}{R(t)} \lim_{\Delta t \rightarrow 0} \left[\frac{1}{\Delta t} P(t < T \leq t + \Delta t) \right] \\ &= \frac{f(t)}{R(t)} = \frac{f(t)}{1 - F(t)} = -\frac{R'(t)}{R(t)}.\end{aligned}$$

D'où,

$$\lambda(t) = -\frac{d \ln [R(t)]}{dt}, \quad (1.2.6)$$

$$\text{et } R(t) = \exp \left[-\int_0^t \lambda(x) dx \right], \quad t > 0. \quad (1.2.7)$$

La courbe en baignoire

La courbe en baignoire donne l'évolution du taux de défaillance $\lambda(t)$ en fonction de l'âge du matériel. Elle comprend trois phases, chacune avec un sens de variation différent (voir la figure 1.5).

Période de jeunesse : La période de jeunesse de l'élément est la période dans laquelle le taux de défaillance décroît rapidement. Elle est caractérisée par des pannes précoces, dites de jeunesse (défauts de conception, défauts de fabrications, etc).

Période de maturité : Est la période de vie utile de l'élément pendant laquelle le taux de défaillance est sensiblement constant. Les défaillances survenant pendant cette période sont dites accidentelles.

Période de vieillesse : Est la période des défaillances d'usure pendant laquelle le taux de défaillance croît rapidement (voir la figure 1.5).

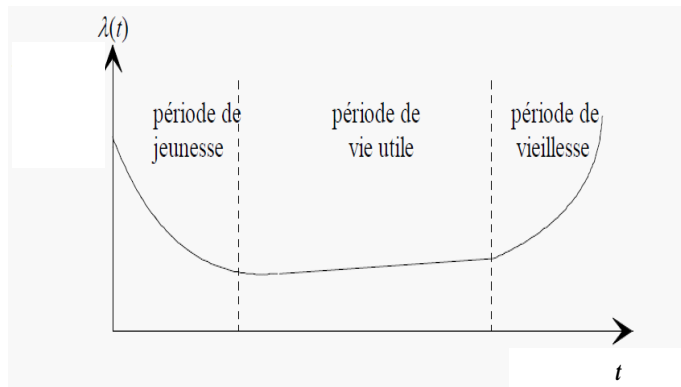


Figure.1.5- Allure de l'évolution du taux de défaillance
« courbe en baignoire ».

Les temps moyen

MTTF (Mean Time To Failure) : Le MTTF se traduit souvent par la durée moyenne de bon fonctionnement. Il est défini comme suit [14] :

$$MTTF = E(T) = \int_0^{+\infty} P(T > t) dt = \int_0^{+\infty} R(t) dt. \quad (1.2.8)$$

MTTR (Mean Time To Repair) : MTTR correspond à la durée moyenne de réparation. Elle n'est pas en général définie que si le matériel est en panne à l'instant initial. Elle est définie comme suit

$$MTTR = \int_0^{+\infty} (1 - M(t)) dt. \quad (1.2.9)$$

MUT (Mean Up Time) : MUT est la durée moyenne de fonctionnement après réparation.

MDT (Mean Down Time) : MDT est la durée moyenne d'indisponibilité des temps d'arrêts.

MTBF(Mean Time Between Failure) : MTBF est la durée moyenne qui sépare deux défaillances successives, comprenant la durée moyenne de bon fonctionnement et la durée moyenne de réparation (voir la figure 1.6).

$$MTBF = MTTF + MTTR. \tag{1.2.10}$$

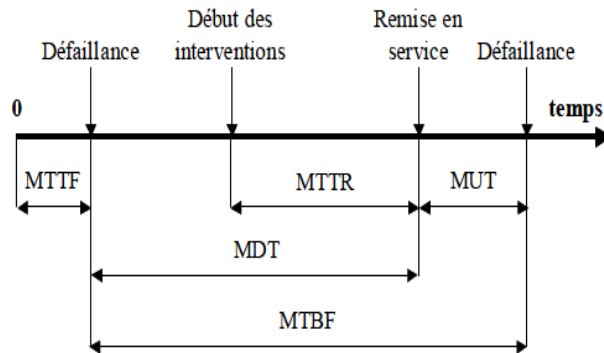


Figure.1.6- Durées moyennes associées à un matériel.

Taux de réparation

Le taux de réparation, noté par $\mu(t)$, est la probabilité pour qu'une entité E , (qui a été en panne pendant un temps t) retrouve son aptitude à remplir sa fonction dans l'unité de temps qui suit. Dans le cas d'une entité non réparable, on procède à un remplacement. Le taux de réparation est défini par [48]:

$$\begin{aligned} \mu(t) &= \lim_{\Delta t \rightarrow 0} \frac{1}{\Delta t} P \left[\begin{array}{l} E \text{ soit réparée (ou remplacée) dans } [t, t + \Delta t] / E \\ \text{est en panne sur } [0, t] \end{array} \right], \forall t \geq 0 \\ &= \lim_{\Delta t \rightarrow 0} P(t < Y \leq t + \Delta t / Y > t). \end{aligned}$$

D'où,

$$\mu(t) = \frac{M'(t)}{1 - M(t)}. \tag{1.2.11}$$

Densité de probabilité de réparation

La densité de probabilité de réparation $m(t)$ est la fonction définie par:

$$m(t) = \frac{dM(t)}{dt} = \mu(t)(1 - M(t)). \quad (1.2.12)$$

1.2.3 Matériel réparable

C'est un matériel qui peut être remis en fonctionnement après avoir subi une défaillance [33]. La vie d'un tel matériel peut être décrite à l'aide de ce chronogramme:

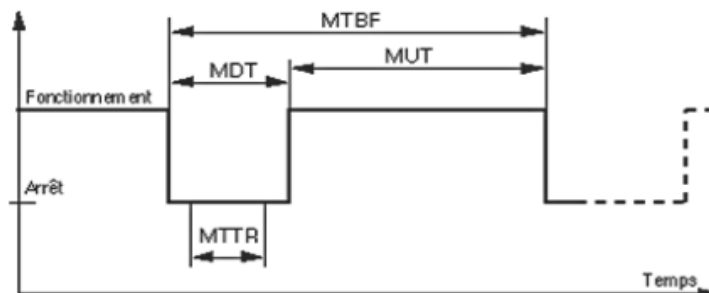


Figure.1.7- Chronogramme d'un équipement réparable.

1.2.4 Matériel irréparable

Le matériel irréparable ne peut pas être remis en fonctionnement. Le chronogramme correspondant est:

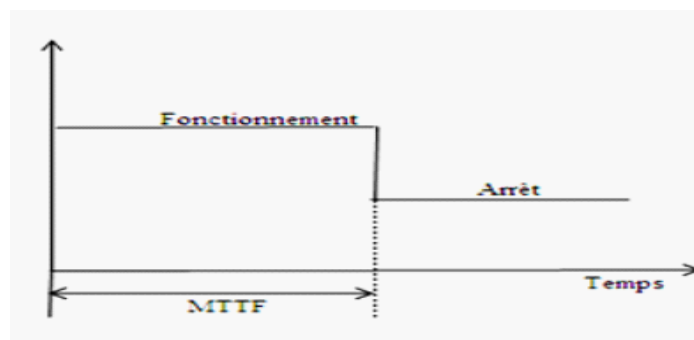


Figure.1.8-Chronogramme d'un équipement irréparable.

1.2.5 Fiabilité des systèmes

Dans l'étude de fiabilité des systèmes réels non réparables, on est amené à décomposer des systèmes en un certain nombre d'éléments ou de sous systèmes. Ces éléments peuvent être indépendants (c'est-à-dire, la défaillance de l'un des éléments du système n'entraîne pas la défaillance d'un autre) ou dépendants [32].

La fiabilité d'un système est déterminée en fonction de la fiabilité de tous ses éléments.

Selon l'incidence de la défaillance de l'un de ces composants sur l'état du système, on distingue trois configurations possibles.

Systeme en série

Ce système fonctionne si tous ses éléments fonctionnent.

Soit un système composé de n équipements supposés indépendants, de durée de vie T_i et de fiabilité R_i , $i = 1, \dots, n$, placés en série (voire la figure 1.9).

Le diagramme de fiabilité du système est le suivant:

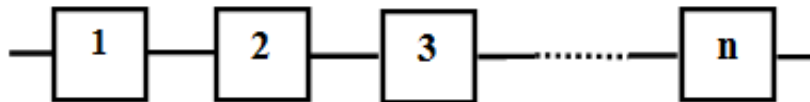


Figure.1.9- Diagramme de fiabilité d'un système en série.

La durée de vie du système T , correspond à

$$T = \min_{i=1,n} T_i.$$

Ce qui signifie qu'il tombe en panne dès que l'un des n éléments tombe en panne. Sa fiabilité est alors celle de l'élément le moins fiable.

Soit $R(t)$ la fiabilité du système. On a alors,

$$R(t) = P(T > t) = P\left(\min_{i=1,n} T_i > t\right) = P(T_1 > t, T_2 > t, \dots, T_n > t). \quad (1.2.13)$$

Les composants étant indépendants, on a

$$R(t) = P(T_1 > t) P(T_2 > t) \dots P(T_n > t) = R_1(t) R_2(t) \dots R_n(t). \quad (1.2.14)$$

D'où,

$$R(t) = \prod_{i=1}^n R_i(t), \quad t > 0. \quad (1.2.15)$$

En utilisant la relation (1.2.15), le taux de défaillance du système, noté $\lambda(t)$ est la somme des taux de défaillance des éléments qui le composent. C'est-à-dire,

$$\lambda(t) = \sum_{i=1}^n \lambda_i(t), \quad t > 0, \quad \text{où } \lambda_i(t) \text{ est le taux de défaillance du } i^{\text{ème}} \text{ composant.} \quad (1.2.16)$$

Remarque 1.2.1 *Le temps moyen de bon fonctionnement du système est inférieur à chacun des temps moyens de fonctionnement de ses composants. C'est-à-dire,*

$$E(T) \leq E(T_i), \quad \forall i = 1, \dots, n.$$

Système en parallèle

Un système de n éléments est dit à configuration parallèle si la panne de tous les éléments est nécessaire pour entraîner la panne du système complet. Autrement dit, Le système est défaillant si et seulement si tous ses éléments sont défaillants. C'est-à-dire que le système est capable de fonctionner avec au moins l'un de ses composants. On distingue deux cas:

Redondance active

Les n composants placés en parallèle fonctionnent en même temps. Le diagramme de fiabilité d'un tel système est donc le suivant:

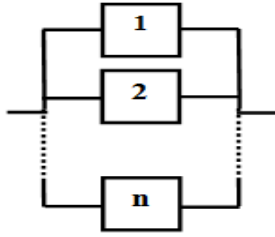


Figure.1.10- Diagramme de fiabilité d'un système en parallèle.

La durée de vie du système en parallèle est le maximum des durées de vie des éléments qui le composent. i.e,

$$T = \max_{i=1,n} T_i. \quad (1.2.17)$$

La fiabilité du système est alors:

$$R(t) = 1 - \prod_{i=1}^n (1 - R_i(t)). \quad (1.2.18)$$

Le temps moyen de bon fonctionnement du système est supérieur au temps de bon fonctionnement de chacun de ses composants. i.e,

$$E(T_i) \leq E(T), \quad \forall i = 1, \dots, n.$$

Redondance passive (ou stand by) Un seul composant fonctionne à la fois. Quand le composant en fonctionnement tombe en panne, il est instantanément remplacé par un des composants en attentes.

Dans ce cas, la durée de vie du système est:

$$T = \sum_{i=1}^n T_i. \quad (1.2.19)$$

Si les n composants sont indépendants et de même loi exponentielle de paramètre λ , $\lambda > 0$, la durée de vie du système en redondance passive T est de loi *gamma* (n, λ) .

Système k parmi n

Le système k parmi n est en état de bon fonctionnement si au moins k composants parmi les n sont en bon fonctionnement (voir la figure 1.11).

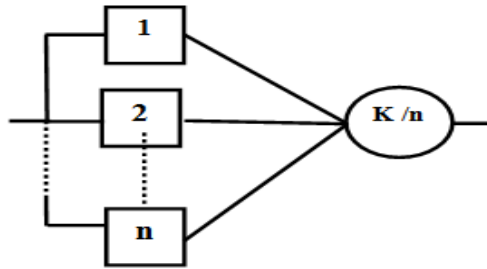


Figure.1.11- Diagramme de fiabilité pour un système k parmi n .

Si en particulier ses éléments ont la même fiabilité $R_1(t)$, le nombre d'éléments non défectueux à l'instant t obéit à une distribution binomiale avec

$$R(t) = \sum_{i=k}^n C_n^i R_1^i(t) [1 - R_1(t)]^{n-i}, \quad t > 0. \quad (1.2.20)$$

Notons que les systèmes k parmi n admettent comme cas particulier les systèmes en série pour ($k = n$) et les systèmes en parallèle pour ($k = 1$).

1.2.6 Principaux modèles de fiabilité

Loi Exponentielle

C'est la loi la plus couramment utilisée en fiabilité, c'est une loi qui décrit la vie des matériels qui subissent des défaillances brutales (aléatoires). Elle s'applique notamment en fiabilité électronique pour décrire la période durant laquelle le taux de défaillance est considéré comme constant, ce qui suppose implicitement que les matériels sont dans leur phase de maturité [41].

Une variable aléatoire T est distribuée suivant une loi exponentielle de paramètre λ , $\lambda > 0$, si sa densité de probabilité f est :

$$f(t) = \lambda \exp(-\lambda t), \quad t \geq 0, \lambda > 0. \quad (1.2.21)$$

- **Fonction de répartition**

$$F(t) = P(T < t) = 1 - \exp(-\lambda t), \quad t \geq 0, \lambda > 0. \quad (1.2.22)$$

- **Fonction de fiabilité**

$$R(t) = P(T > t) = 1 - F(t) = \exp(-\lambda t), \quad t \geq 0, \lambda > 0. \quad (1.2.23)$$

- **Taux de défaillance**

$$\lambda(t) = \frac{f(t)}{R(t)} = \frac{\lambda \exp(-\lambda t)}{\exp(-\lambda t)} = \lambda, \quad \forall t. \quad (1.2.24)$$

- **MTTF**

$$E(T) = \frac{1}{\lambda}. \quad (1.2.25)$$

- **La variance**

$$V(T) = \frac{1}{\lambda^2}. \quad (1.2.26)$$

La représentation graphique de la fonction de densité de probabilité $f(t)$ de la loi exponentielle est donnée par la figure ci-dessous.

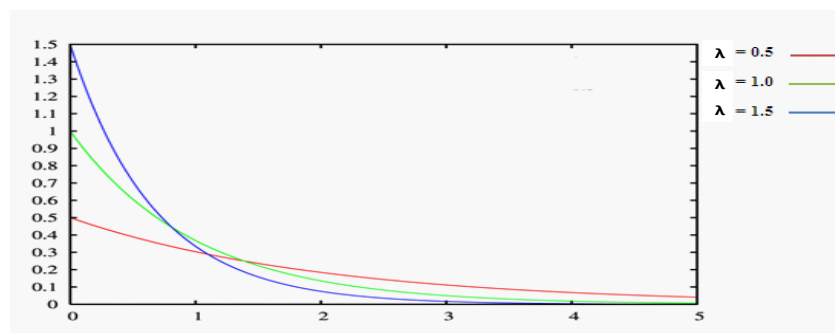


Figure.1.12- Densité de probabilité de la loi exponentielle en fonction de λ .

Propriété sans mémoire

Une propriété principale de la loi exponentielle est d'être sans mémoire (Memoryless property, en anglais). C'est-à-dire,

$$P(T \geq t + dt / T \geq t) = \frac{\exp(-\lambda(t + dt))}{\exp(-\lambda t)} = \exp(-\lambda dt), \quad t \geq 0, dt \geq 0. \quad (1.2.27)$$

Le résultat donné par la formule ci-dessus montre que la loi conditionnelle de la durée de vie d'un dispositif qui a fonctionné sans tomber en panne jusqu'à l'instant t est identique à la loi de la durée de vie d'un nouveau dispositif. Ceci signifie qu'à l'instant t , le dispositif est considéré comme neuf ("as good as new", en anglais), de durée de vie exponentielle de paramètre λ [47].

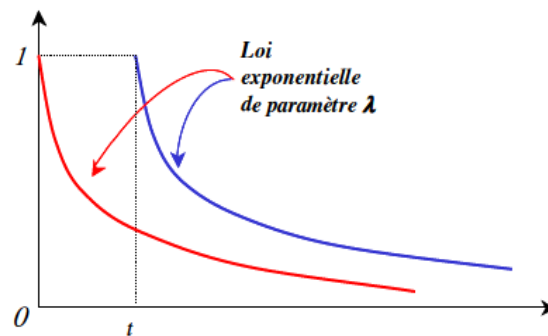


Figure.1.13- Propriété sans mémoire de la loi exponentielle.

Remarques 1.2.2 1) *La loi exponentielle est la seule loi continue ayant un taux de défaillance constant, ce qui signifie que cette loi modélise les durées de vie des systèmes qui ne s'usent pas, et qui ne s'améliore pas (la phase utile). Sa propriété simplifie énormément les calculs, et les résultats obtenus sont faciles à interpréter.*

2) *La distribution exponentielle est associée aux processus de Poisson que l'on définit comme un processus qui génère des événements, dont les temps d'occurrence sont indépendants et distribués identiquement selon une loi exponentielle.*

3) *L'espérance de T qui correspond à la durée de vie moyenne d'un matériel est plus accessible que les experts peuvent plus facilement l'estimer par la méthode classique que par la démarche bayésienne.*

Loi Weibull

Weibull, mathématicien suédois a créé la distribution Weibull qui porte son nom, qui a été utilisée pour la première fois pour décrire la fatigue des équipements mécaniques en gestion de maintenance parce qu'elle correspond à un modèle bien adapté à l'étude statistique des défaillances [30]. Cette loi est très intéressante car elle recouvre un grand nombre de distributions expérimentales continues généralement appliquées aux durées de vie de nombreux types de matériels, et encore a l'avantage d'être très souple et de pouvoir représenter les différentes phases de la vie d'un matériel (jeunesse, vie utile et vieillissement). La loi Weibull n'est applicable en toute rigueur qu'à des composants non réparables et seulement aux premières pannes des matériels réparables.

La loi de Weibull, notée $W(\gamma, \eta, \beta)$, est une loi continue et repose sur trois paramètres qui sont (voir la figure 1.14):

- γ : appelé paramètre de position qui représente le décalage qui existe entre le début de l'observation et le début du processus examiné.
- η : appelé paramètre d'échelle, lié à la durée de vie caractéristique, c'est en quelque sorte une approximation de la durée de vie du matériel considéré.
- β : appelé paramètre de forme, caractérise la cinétique d'évolution des défaillances.

Suivant les différentes valeurs de β , le taux de défaillance est [11]:

- pour $\beta < 1$, décroissant avec le temps, ce qui est représentatif de la période de jeunesse des matériels ;
- pour $\beta = 1$, constant avec le temps, ceci correspond à la période de maturité, où on retrouve la loi exponentielle ;
- pour $\beta > 1$, croissant avec le temps, on aborde la période de vieillissement ou d'usure. C'est là qu'il faudra intervenir de façon préventive.

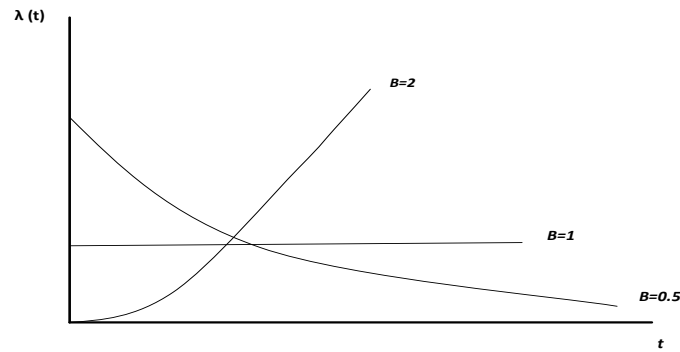


Figure.1.14- Le taux de défaillance d'une Weibull pour différentes valeurs de β .

Soit une variable aléatoire T distribuée selon la $W(\gamma, \eta, \beta)$. La densité de probabilité f est donnée par:

$$f(t) = \frac{\beta (t - \gamma)^{\beta-1}}{\eta^\beta} \exp \left[- \left(\frac{t - \gamma}{\eta} \right)^\beta \right], \beta > 0, \eta > 0, t > \gamma. \quad (1.2.28)$$

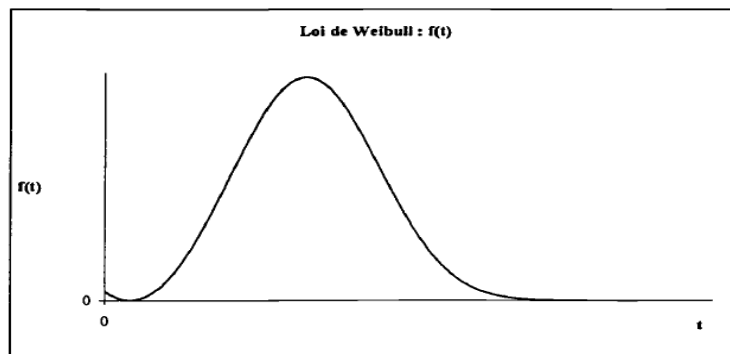


Figure.1.15- Densité de probabilité de la loi Weibull.

- **Fonction de répartition**

$$F(t) = 1 - \exp \left[- \left(\frac{t - \gamma}{\eta} \right)^\beta \right], t > \gamma. \quad (1.2.29)$$

- **Fonction de fiabilité**

$$R(t) = 1 - F(t) = \exp \left[- \left(\frac{t - \gamma}{\eta} \right)^\beta \right], \quad t > \gamma. \quad (1.2.30)$$

Par ailleurs, les caractéristiques de la variable aléatoire T sont données par,

• **Taux de défaillance**

$$\lambda(t) = \frac{\beta}{\eta} \left(\frac{t - \gamma}{\eta} \right)^{\beta-1}, \quad t > \gamma. \quad (1.2.31)$$

• **MTTF**

$$E(T) = \gamma + \eta \Gamma \left(\frac{1 + \beta}{\beta} \right), \quad (1.2.32)$$

où, Γ est la fonction gamma d'Euler définie par

$$\Gamma(\alpha) = \begin{cases} \int_0^{+\infty} t^{\alpha-1} e^{-t} dt & \text{si } \alpha \in \mathbb{R} \\ (\alpha - 1)! & \text{si } \alpha \in \mathbb{N}^*. \end{cases} \quad (1.2.33)$$

• **La variance**

$$V(T) = \eta^2 \left[\Gamma \left(\frac{2}{\beta} + 1 \right) - \Gamma^2 \left(1 + \frac{1}{\beta} \right) \right]. \quad (1.2.34)$$

Remarques 1.2.3 Pour des valeurs particulières des paramètres, on obtient des lois connues:

- Si $\gamma = 0$ et $\beta = 1$, on obtient la loi exponentielle de paramètre $\left(\frac{1}{\eta} \right)$.

Avec

$$\lambda(t) = \frac{1}{\eta} = \frac{1}{MTTF} = cst \quad \forall t > 0. \quad (1.2.35)$$

- Si $\gamma = 0$ et $\beta \simeq 3.5$, on est dans le cas d'une distribution normale, car la loi normale modélise la vie des matériels qui subissent à une vieillissement forte [29].

Loi Gamma

La loi gamma est généralement une loi à deux paramètres positifs: α le paramètre de forme, équivalent à un nombre de défaillances, et β le paramètre d'échelle, correspondant à un temps de fonctionnement. Elle représente la loi de probabilité d'occurrence de α

événements dans un processus poissonnien. Cette loi est souvent utilisée pour modéliser les temps de défaillances d'un matériel. Par exemple si T_i est le temps entre les défaillances successives d'un matériel et distribuée selon une loi exponentielle de paramètre β , alors le temps cumulé d'apparition de α défaillances suit une loi gamma(α, β) [35].

- **Densité de probabilité**

$$f(t) = \frac{\beta^\alpha}{\Gamma(\alpha)} t^{\alpha-1} \exp(-\beta t), \quad t > 0. \quad (1.2.36)$$

- **Fonction de répartition**

$$F(t) = \int_0^t \frac{\beta^\alpha}{\Gamma(\alpha)} t^{\alpha-1} \exp(-\beta t) dx, \quad t > 0, \quad (1.2.37)$$

où

$$\Gamma(\beta) = \int_0^\infty t^{\beta-1} \exp(-\beta t) dt. \quad (1.2.38)$$

- **Fonction de fiabilité**

$$R(t) = 1 - F(t) = \int_t^\infty \frac{\beta^\alpha}{\Gamma(\alpha)} t^{\alpha-1} \exp(-\beta t) dx, \quad t > 0. \quad (1.2.39)$$

- **Taux de défaillance**

$$\lambda(t) = \frac{\beta^\alpha t^{\alpha-1} \exp(-t\beta)}{\int_t^\infty \Gamma(\beta) f(x) dx}, \quad t > 0. \quad (1.2.40)$$

- **MTTF**

$$E(T) = \frac{\alpha}{\beta}. \quad (1.2.41)$$

- **La variance**

$$V(T) = \frac{\alpha}{\beta^2}. \quad (1.2.42)$$

Remarques 1.2.4 La loi gamma s'identifie à d'autres lois. Ainsi,

- Si $T \sim \text{gamma}(\alpha, \beta)$, alors $\frac{1}{T} \sim \text{Gamma Inverse}\left(\alpha, \frac{1}{\beta}\right)$.
- Si $\alpha=1$, on retrouve la loi exponentielle de paramètre β .

- Si β est un entier, on obtient la loi d'Erlang qui a pour densité de probabilité :

$$f(t) = \frac{\beta^n t^{n-1}}{(n-1)!} \exp(-\beta t), \quad t > 0. \quad (1.2.43)$$

- Et pour $n = 1$, on retrouve encore une loi exponentielle de paramètre β .

- Si la variable T a pour densité une Gamma $(1, \beta)$, alors $T = X^\alpha$ suit une loi Weibull (α, β) .

- Pour $\alpha = \frac{1}{2}$ et $\beta = \frac{v}{2}$, avec v est un entier, on obtient la loi du χ_v^2 (Khi-Deux) à v degrés de liberté.

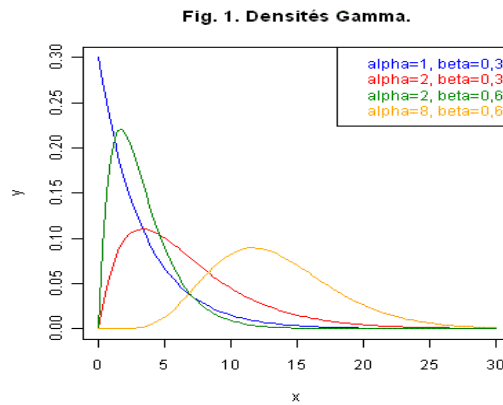


Figure.1.16- Densité de probabilité de la loi gamma pour différentes valeurs de α et β .

Loi Bêta

La distribution Bêta peut prendre toutes les formes possibles et peut donc représenter de nombreux phénomènes. Elle est fréquemment utilisée dans la démarche bayésienne et en contrôle de qualité. C'est la loi d'une variable aléatoire T à support dans $[0, 1]$. Elle dépend de deux paramètres a et b . On la note $\beta e(a, b)$ [41].

- **Densité de probabilité:**

$$f(t) = \frac{1}{B(a, b)} t^{a-1} (1-t)^{b-1}, \quad \text{avec } t \in [0, 1] \text{ et } a, b > 0, \quad (1.2.44)$$

avec,

$$B(a, b) = \left(\frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} \right)^{-1}. \quad (1.2.45)$$

a et b sont appelés les paramètres de forme de la loi. Ils répondent à la relation de symétrie. C'est-à-dire:

$$f(t; a, b) = f(1-t; b, a).$$

- **Fonction de répartition**

$$F(t) = \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} \int_0^t x^{a-1} (1-x)^{b-1} dx, \quad t \in \mathbb{R}. \quad (1.2.46)$$

- **L'espérance**

$$E(T) = \frac{a}{a+b}. \quad (1.2.47)$$

- **La variance**

$$V(T) = \frac{ab}{(a+b+1)(a+b)^2}. \quad (1.2.48)$$

Remarque 1.2.5 Cette loi est souvent utilisée en statistique bayésienne (survie d'un matériel) comme distribution a priori de la probabilité d'un évènement qui suit une distribution binomiale. Les deux lois étant conjuguées naturelles, le calcul de la distribution a posteriori est alors simplifié. Elle sert à exprimer en particulier la fiabilité $R(t)$ d'un système composé de k matériels.

Loi Binomiale

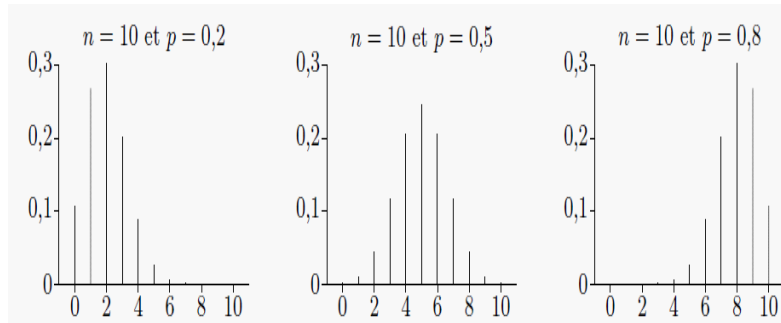
Dans beaucoup d'applications d'ingénierie, les évènements constitués d'une répétition d'essais peuvent être formulés comme suit : occurrence ou non-occurrence, succès ou échec, bon ou mauvais,...etc [7].

Soit p la probabilité de réalisation d'un évènement A . La variable aléatoire discrète T , représentant le nombre de réalisations de l'évènement A au cours de n expériences, est distribuée suivant une loi binomiale de paramètres (n, p) tel que

$$P[T = k] = C_n^k p^k (1-p)^{n-k}, \quad k \in \{0, 1, \dots, n\}, \quad 0 \leq p \leq 1. \quad (1.2.49)$$

Avec,

$$C_n^k = \frac{n!}{k!(n-k)!}. \quad (1.2.50)$$

Figure.1.17- La loi Binomiale (n, p) .

- **Fonction de répartition**

$$F(k) = P[T \leq k] = \sum_{i=0}^k C_n^i p^i (1-p)^{n-i}, \quad k \in \{0, 1, \dots, n\}. \quad (1.2.51)$$

- **MTTF**

$$E(T) = np. \quad (1.2.52)$$

- **La variance**

$$V(T) = np(1-p). \quad (1.2.53)$$

Remarque 1.2.6 On rencontre cette distribution lorsqu'un matériel a une probabilité γ de défaillance à la sollicitation. Au bout de n sollicitations, le nombre de défaillances est distribué suivant une loi Binomiale (n, p) .

Propriétés

- Lorsque $n \rightarrow \infty$ et si np reste constant, la loi Binomiale tend vers une loi de Poisson de paramètre $\lambda = np$;
- Lorsque $n \rightarrow \infty$ la loi Binomiale tend vers une loi normale de moyenne np et de variance $np(1-p)$.

Remarque 1.2.7 D'autres modèles de fiabilité peuvent être utilisés tel que: loi Normale, loi log-Normale, Loi de Pareto et loi de Poisson (voir annexe A)

Conclusion

L'estimation des paramètres de fiabilité des systèmes fait souvent l'objet de développements théoriques et pratiques importants. Les méthodes Bayésiennes permettent d'évaluer la distribution de probabilité de ces paramètres, à partir de la connaissance a priori que l'on possède sur le système (expériences, essais effectués dans le passé, comportement de matériels similaires) ou par le retour d'expérience.

La mise en commun des données par l'approche Bayésienne a comme intérêt:

- d'utiliser le maximum d'informations disponibles ;
- de valider un objectif de fiabilité avec un niveau de confiance ;
- de permettre l'actualisation du savoir automatiquement ;
- de se prononcer dans le cas où aucune défaillance n'a été constatée ;
- de trouver une valeur (taux de défaillance) la plus pertinente.

Modélisation et choix de la loi a priori

Le point le plus critiqué de l'analyse Bayésienne est le choix de la loi a priori qui représente la clé de voute de l'inférence Bayésienne, et sa détermination est donc l'étape la plus importante dans la mise en œuvre de cette inférence.

Evidement dans la pratique ce choix est rendu d'autant plus difficile lorsque l'on ne dispose pas d'informations suffisamment précises pour conduire à une détermination exacte de la loi a priori au sens où plusieurs lois de probabilité peuvent être compatibles avec cette information. Parfois le statisticien n'a pas le temps ni les ressources pour construire une loi a priori exacte et il doit compléter l'information partielle qu'il a rassemblé à l'aide des données subjectives afin d'obtenir une loi a priori.

Même dans le cas où l'information a priori est disponible, il est rare de pouvoir proposer une détermination exacte de la loi a priori à partir de cette information. Il est donc nécessaire de faire un choix arbitraire de la loi a priori, ce qui peut avoir un impact considérable sur l'inférence qui en découle. En particulier dans le cas paramétrique, plusieurs lois peuvent être candidates pour la modélisation, sachant que le choix d'une certaine loi peu conduire à des résultats a postériori différents [43].

Une difficulté majeure pour le choix de la loi a priori, apparaît aussi lorsque l'espace des paramètres Θ est non borné ou non dénombrable (cas des lois a priori impropres) [43]. Pour ces raisons, et à moins de connaître la distribution liée au paramètre, il est souvent difficile de sélectionner une loi a priori.

Nous considérerons deux cas : l'approche a priori conjuguée qui nécessite une quantité limitée d'informations (cas informatif) et l'approche non informative qui est obtenue à partir de la distribution de l'échantillon.

2.1 Lois a priori non informatives

La loi a priori non informative représente une ignorance sur le problème considéré mais ne signifie pas que l'on sache absolument rien sur la distribution statistique du paramètre θ . En effet, on connaît au moins son domaine de variation. Ces lois doivent être particulièrement construites à partir de la distribution de l'échantillonnage puisque c'est le seul moyen disponible pour avoir des informations sur le paramètre θ . À cet égard, les lois a priori non informatives peuvent être considérées comme des lois de références, auxquelles chacun pourrait avoir recours quand toute information a priori sur θ est absente [10].

Nous décrirons, dans ce qui suit, quelques unes des techniques les plus populaires dans la construction des lois a priori non informatives.

2.1.1 Loi a priori de Laplace

Historiquement, Laplace fut le premier à utiliser les techniques non informatives. Le nombre de cartes blanches et noires dans une urne, et la proportion de naissances mâles à Paris, sont des exemples de références.

Le raisonnement de Laplace se fondait sur l'équiprobabilité des événements élémentaires, et donc de mener la loi uniforme à la loi a priori du paramètre [3].

Exemple 2.1.1 (*Laplace, 1773. les cartes blanches et noires*)[43]

Une urne contient un nombre n de cartes noires et blanches. Si la première carte sortie de l'urne est blanche, quelle est la probabilité que la proportion θ de cartes blanches soit θ_0 ? Pour résoudre ce problème, Laplace suppose que tous les nombres de 2 à $n - 1$ sont équiprobables comme valeurs de θn , donc θ soit uniformément distribuée sur $\{2/n, \dots, (n - 1)/n\}$.

La distribution a posteriori de θ peut être alors calculée en utilisant le théorème de Bayes,

$$\pi(\theta = \theta_0 | \text{données}) = \frac{\theta_0 \times 1/(n-2)}{\sum_{\theta=2/n}^{(n-1/n)} \theta \times 1/(n-2)} = \frac{n\theta_0}{n(n-1)/2 - 1}.$$

Evidemment le choix précédent de la distribution a priori peut être contesté comme étant partiellement arbitraire. Cependant, dans la perspective de la théorie des probabilités de Laplace, la plupart des événements peuvent être décomposés en événements équiprobables élémentaires et par conséquent, dans ce cas particulier, il semble raisonnable de considérer les événements $\{\theta = i/n\}$, $i = 2, \dots, n-1$, comme élémentaires. Un raisonnement similaire justifie l'exemple suivant.

Exemple 2.1.2 (Laplace, 1786. La naissance mâles à Paris)

Considérons la proportion de naissances masculines à Paris. Laplace veut vérifier que la probabilité θ d'une naissance masculine dépasse $1/2$. En observant 251527 naissances masculines et 241945 naissances féminines en 1785 et en supposant que θ a pour distribution a priori la loi uniforme sur $[0, 1]$. Laplace obtient

$$p(\theta \leq 1/2 | (251527; 241945)) = 1.15 \times 10^{-42}.$$

(Voir Stigler, 1986, p. 134 et l'Exercice 1.8). Il déduit alors que la probabilité θ est très vraisemblablement supérieure à 50%. Utilisant de nouveau une distribution a priori uniforme, il compare aussi les naissances masculines à Londres et à Paris et en déduit que la probabilité d'une naissance masculine est aussi significativement supérieure à 50% en Angleterre.

Remarques 2.1.1 1) Les lois résultantes sont impropres quand l'espace des paramètres n'est pas compact.

2) Le principe des événements équiprobables de Laplace n'est pas cohérent en termes de partitionnement : si $\Theta = \{\theta_1, \theta_2\}$ la règle de Laplace donne $\pi(\theta_1) = \pi(\theta_2) = 1/2$ mais, si la définition de Θ est plus détaillée, avec $\Theta = \{\theta_1, \theta_2, \theta_3\}$ la règle de Laplace mène à $\pi(\theta_1) = 1/3$. Ce qui n'est pas évidemment cohérent avec la première formulation.

3) Le problème le plus fondamentale, concerne le problème de l'invariance par reparamétrisation. Si on passe de θ à $\eta = g(\theta)$, avec $\theta \in \Theta$ par une transformation bijective g ,

l'information a priori reste totalement inexistante et ne devrait pas être modifiée. Cependant, si $\pi(\theta) = 1$ la loi a priori sur η est :

$$\pi^*(\eta) = \left| \frac{d}{d\eta} g^{-1}(\eta) \right|. \quad (2.1.1)$$

Donc $\pi(\eta)$ est le plus souvent non constante.

Exemple 2.1.3 (*Shafer, pp. 23-24*)[25]

Shafer a donné un simple exemple pour comprendre mieux le deuxième point de la remarque.

Soit $\Theta = \{\theta_1, \theta_2\}$, où θ_1 et θ_2 sont des événements définis par :

θ_1 : *"il y ait une vie en orbite autour l'étoile Sirius."*

θ_2 : *"il n y ait aucune."*

La règle de Laplace donne:

$$P(\{\theta_1\}) = P(\{\theta_2\}) = \frac{1}{2}.$$

Soit maintenant $\Omega = \{\omega_1, \omega_2, \omega_3\}$, où $\omega_i, i = 1, \dots, 3$ sont les évts définis par :

ω_1 : *" il y ait une vie en orbite autour l'étoile Sirius."*

ω_2 : *" il ya des planètes mais il n y ait pas de vie."*

ω_3 : *" il n y ait pas des planètes."*

La règle de Laplace donne :

$$P(\{\omega_1\}) = P(\{\omega_2\}) = P(\{\omega_3\}) = \frac{1}{3}.$$

Le paradoxe est que la probabilité de l'évènement qu'il y a une vie, $P(\{\theta_1\}) = \frac{1}{2}$ si nous adoptons la première formulation, et $P(\{\omega_1\}) = \frac{1}{3}$, si nous adoptons la deuxième formulation. D'où les deux formules ne sont pas cohérentes.

Exemple 2.1.4 *Si θ est la proportion de naissances mâles et suit une loi Uniforme sur $[0, 1]$, alors le paramètre de rapport des chances $\eta = \frac{\theta}{1-\theta}$ suit une loi a priori de densité $\frac{1}{(1+\eta)^2}$, qui est donc non constante. En effet,*

on a:

$$\eta = \frac{\theta}{1-\theta} \implies g^{-1}(\eta) = \theta \implies g^{-1}(\eta) = \frac{\eta}{\eta+1}$$

et

$$\left| \frac{d}{d\eta} g^{-1}(\eta) \right| = \frac{1}{(\eta + 1)^2}.$$

D'où,

$$\pi^*(\eta) = \left| \frac{d}{d\eta} g^{-1}(\eta) \right| = \frac{1}{(\eta + 1)^2},$$

qui n'est plus une constante.

2.1.2 Loi a priori de Jeffreys

Jeffreys, en 1961 propose une méthode de construction de la loi a priori non informative, en utilisant l'information de Fisher $I_n(\theta)$ qui représente une mesure de la quantité d'information sur θ obtenue par les observations [26].

Définition 2.1.1 Soit θ un paramètre réel. On appelle une loi a priori non informative de Jeffreys la loi (éventuellement impropre) de densité

$$\pi_J(\theta) \propto [I_n(\theta)]^{1/2} 1_{\Theta}(\theta), \quad (2.1.2)$$

où $I_n(\theta)$, est l'information de Fisher apportée par $\underline{X} = (X_1, \dots, X_n)$ sur θ , définie par

$$I_n(\theta) = \left[\left(\frac{\partial}{\partial \theta} \log L(\theta, \underline{x}) \right)^2 \right] = -E \left[\frac{\partial^2}{\partial \theta^2} \log L(\theta, \underline{x}) \right], \text{ si le domaine de } \underline{X} \text{ ne dépend pas de } \theta. \quad (2.1.3)$$

Remarques 2.1.2 1) Si on dispose d'une seule observation, l'information de Fisher s'écrit sous la forme

$$I_X(\theta) = -E \left[\frac{\partial^2}{\partial \theta^2} \log f(x|\theta) \right], \text{ si le domaine de } X \text{ ne dépend pas de } \theta. \quad (2.1.4)$$

2) Plus $I_n(\theta)$ est grande plus, l'observation apporte de l'information sur θ . Il est donc naturel de favoriser les valeurs de θ pour les quelles $I_n(\theta)$ est grande au sens rendre plus probable suivant $\pi_J(\theta)$.

3) Le choix de la loi a priori dépendant de l'information de Fisher se justifie par le fait que $I_n(\theta)$ est largement acceptée comme un indicateur de la quantité d'information apportée par l'observation \underline{x} sur θ (voir Fisher 1956) [3].

4) La loi a priori de Jeffreys est invariante par reparamétrisation. Puisque pour une transformation bijective donnée h qui transforme le paramètre θ en $\lambda = h(\theta)$. On peut définir un nouveau modèle statistique d'échantillon \underline{x} et de vraisemblance, tel que [18],

$$\tilde{L}(\underline{x}|\lambda) = L(\underline{x}|h^{-1}(\lambda))$$

On note

$$l(\underline{x}, \lambda) = \log L(\underline{x}, \lambda).$$

La loi a priori de Jeffreys de λ est:

$$\tilde{\pi}_J(\lambda) \propto \left| E \left(l'(\underline{x}|h^{-1}(\lambda))'' \right) \right|^{\frac{1}{2}}$$

$$\begin{aligned} \text{et } l(\underline{x}|h^{-1}(\lambda))'' &= \left(l'(\underline{x}, h^{-1}(\lambda)) \left(h^{-1}(\lambda)' \right) \right)' \\ &= l''(\underline{x}|h^{-1}(\lambda)) \left(h^{-1}(\lambda)' \right) \left(h^{-1}(\lambda)' \right) + l'(\underline{x}|h^{-1}(\lambda)) \left(h^{-1}(\lambda)'' \right) \end{aligned}$$

D'où,

$$\begin{aligned} \tilde{\pi}_J(\lambda) &\propto \left| E \left(l''(\underline{x}|h^{-1}(\lambda)) \left(h^{-1}(\lambda)' \right)^2 + l'(\underline{x}|h^{-1}(\lambda)) \left(h^{-1}(\lambda)'' \right) \right) \right|^{\frac{1}{2}} \\ &\propto \left| E \left(l''(\underline{x}|h^{-1}(\lambda)) \left(h^{-1}(\lambda)' \right)^2 \right) \right|^{\frac{1}{2}}, \end{aligned}$$

car l'espérance de la dérivée première de log vraisemblance est égale à 0.

$$\begin{aligned} \Rightarrow \tilde{\pi}_J(\lambda) &\propto \left| E \left(l''(\underline{x}, h^{-1}(\lambda)) \right) \right|^{\frac{1}{2}} \left| h^{-1}(\lambda)' \right| \\ \Rightarrow \tilde{\pi}_J(\lambda) &\propto \pi_J(h^{-1}(\lambda)) \left| h^{-1}(\lambda)' \right|. \end{aligned}$$

5) Dans le cadre de la théorie du maximum de vraisemblance, le déterminant de la quantité d'information de Fisher représente ce que l'on appelle la variance généralisée.

6) L'approche de Jeffreys fournit une meilleure technique pour obtenir une loi a priori non

informatives, de plus elle nous permet de retrouver les estimateurs classiques. Mais elle a été critiquée par certains Bayésiens comme étant un outil sans justification subjective en terme d'information a priori.

7) Dans le cas multidimensionnel $\theta = (\theta_1, \dots, \theta_q)$, $\theta \in \mathbb{R}^q$, la loi a priori non informative de Jeffreys s'écrit

$$\pi_J(\theta) \propto [\det(I_n(\theta))]^{1/2}. \quad (2.1.5)$$

où $I_n(\theta)$ la matrice d'information de Fisher, donnée par les éléments

$$I_n(\theta)|_{i,j} = -E \left[\frac{\partial^2}{\partial \theta_i \partial \theta_j} \log L(\theta, \underline{x}) \right], \quad (i, j = 1, \dots, q) \quad (2.1.6)$$

Exemple 2.1.5 (Cas unidimensionnel)

Soit X suit une loi exponentielle de paramètre θ , $\theta > 0$

$$f(x|\theta) = \theta \exp(-\theta x), \quad x > 0$$

On calcule l'information de Fisher apportée par X sur θ

$$\Rightarrow \log(f(x|\theta)) = \text{Log}\theta - \theta x,$$

On a:

$$\frac{\partial}{\partial \theta} \log f(x|\theta) = \frac{1}{\theta} - x,$$

$$\frac{\partial^2}{\partial \theta^2} \log f(x|\theta) = -\frac{1}{\theta^2},$$

d'où,

$$I_X(\theta) = -E \left[\frac{\partial^2}{\partial \theta^2} \log f(x|\theta) \right] = -E \left(-\frac{1}{\theta^2} \right) = \frac{1}{\theta^2}.$$

La loi a priori non informative de Jeffreys est alors,

$$\pi_J(\theta) \propto [I(\theta)]^{1/2} = \frac{1}{\theta}, \quad \theta > 0.$$

Ce type de construction conduit très souvent à des lois a priori impropres.

Exemple 2.1.6 (Berger et Yang 1995, cas multidimensionnel)

Soit $X \sim N(\mu, \sigma^2)$, où μ est la moyenne et σ^2 la variance de X

On suppose que μ et σ^2 sont inconnus.

La fonction de densité de X s'écrit

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2\right), \quad x \in \mathbb{R}.$$

On note $\theta = (\mu, \sigma^2)$.

$$\begin{aligned} f(x|\theta) &\propto \frac{1}{\sigma} \exp\left\{-\frac{(x-\mu)^2}{2\sigma^2}\right\} \\ \Rightarrow \log f(x|\theta) &\propto -\frac{1}{2} \log \sigma^2 - \frac{(x-\mu)^2}{2\sigma^2} \end{aligned}$$

On a:

$$\begin{aligned} \frac{\partial}{\partial \mu} \log f(x|\theta) &= \frac{(x-\mu)}{\sigma^2}, \\ \frac{\partial^2}{\partial \mu^2} \log f(x|\theta) &= -\frac{1}{\sigma^2}, \\ \frac{\partial^2}{\partial \mu \partial \sigma^2} \log f(x|\theta) &= -\frac{(x-\mu)}{2(\sigma^2)^2}, \\ \frac{\partial}{\partial \sigma^2} \log f(x|\theta) &= -\frac{1}{2\sigma^2} + \frac{(x-\mu)^2}{2(\sigma^2)^2}, \\ \frac{\partial^2}{\partial (\sigma^2)^2} \log f(x|\theta) &= \frac{1}{2(\sigma^2)^2} - \frac{(x-\mu)^2}{(\sigma^2)^3}, \end{aligned}$$

La matrice d'information de Fisher s'obtient en calculant, l'espérance mathématique des dérivées secondes en utilisant le fait que $E(X) = \mu$ et $E(X - \mu)^2 = \sigma^2$

On obtient

$$I_n(\theta) = \begin{pmatrix} \frac{1}{\sigma^2} & 0 \\ 0 & \frac{1}{2(\sigma^2)^2} \end{pmatrix} \Rightarrow \det [I_n(\theta)] = \frac{1}{2\sigma^6}$$

D'où, la loi a priori de Jeffreys est

$$\pi(\theta) \propto \frac{1}{\sigma^3}, \quad \sigma > 0, \quad \mu \in \mathbb{R}.$$

2.1.3 Loi a priori invariante

La loi a priori invariante est obtenue lorsque seuls les paramètres de la distribution de θ changent à partir d'une transformation de θ .

Principe de la loi a priori invariante

Il s'agit de trouver une distribution a priori invariante par certaines opérations. Si la famille de distributions est stable par une classe de transformations, on pourra chercher une distribution a priori invariante par cette famille de transformations [49].

1. Invariance par translation

Soit X est distribuée selon une densité f appartenant à une famille de lois de la forme $f(x|\theta) = f(x - \theta)$, où θ est un paramètre de position inconnu. Dans ce cas, f est invariante par translation.

C'est-à-dire, que si $x \sim f(x|\theta)$ alors $Y = X + a$ est distribuée selon $f(y - \beta)$, avec $\beta = \theta + a$ un paramètre de position inconnu et a est une constante.

X et Y ont une distribution appartenant à la même famille de lois, mais avec un paramètre de position différent.

La loi a priori pour θ doit être invariante par translation donc

$$\begin{aligned}\pi(\theta) &= \pi(\theta - a) \quad \forall a \\ &\Rightarrow \pi(\theta) = 1.\end{aligned}\tag{2.1.7}$$

La loi a priori non informative pour le paramètre de position est la loi uniforme sur Θ .

Remarques 2.1.3 1) *Les paramètres de position caractérisent l'ordre de grandeur des observations. Ils permettent aussi de situer des distributions les unes par rapport aux autres, les paramètres de positions les plus utilisés sont: le mode, médiane et la moyenne.*
2) *Un autre argument conduisant au même résultat est le fait que θ et β sont des paramètres de position dans le même modèle, ils devraient avoir une même loi a priori.*

2. Invariance par changement d'échelle

Soit X est distribuée selon $\frac{1}{\sigma}f(\frac{x}{\sigma})$, avec $\sigma > 0$ un paramètre d'échelle, f est alors invariante par le changement d'échelle. La variable $Y = \frac{X}{\sigma}$ a pour densité la fonction f .

X et Y ont une distribution de la même famille mais avec un paramètre d'échelle différent. Le fait que f est paramétrée par un paramètre d'échelle, la loi a priori doit être

invariante par un changement d'échelle:

$$\begin{aligned}\pi(A) &= \pi(A/c) \quad \forall A \in (0, +\infty), \quad c > 0 \\ \Rightarrow \pi(\sigma) &= \frac{1}{c} \pi\left(\frac{\sigma}{c}\right), \quad c > 0 \\ \Rightarrow \pi(\sigma) &= \sigma^{-1}, \quad \sigma > 0.\end{aligned}\tag{2.1.8}$$

D'où, la loi a priori invariante n'est plus une constante.

L'approche invariante n'est que partiellement satisfaisante, car elle implique la référence à une structure d'invariance qui peut être parfois choisie de plusieurs manières, ne pas exister, ou être sans intérêt pour le décideur.

Deux cas usuels sont déduits afin de déterminer la loi apriori invariante, à partir de l'exemple précédent.

- Si X est distribuée selon une densité f qui est invariante par translation, c'est-à-dire f appartient à une famille de lois de la forme $f(x - \theta)$, où θ est un paramètre de position inconnu alors, $\pi(\theta) \propto 1$. La loi a priori est alors la loi uniforme sur Θ .
- Si X est distribuée selon une densité f qui est invariante par changement d'échelle, c'est-à-dire f appartient à une famille de lois de la forme $\frac{1}{\theta} f\left(\frac{x}{\theta}\right)$, où $\theta > 0$ est un paramètre d'échelle inconnu. Alors, $\pi(\theta) \propto \theta^{-1}$.

Exemple 2.1.7 Application à la loi Normale

Considérons une famille de lois stables par les transformations affines.

Soit X une v.a de densité $f_X(x)$ associée à la loi Normale $N(\mu, \sigma^2)$, alors f_X appartient à la famille de lois stables par transformations affines.

La v.a Y , définie par : $Y = \sigma X + \mu$, a pour densité la fonction

$$f_Y(y) = \sigma^{-1} f_X\left(\frac{y - \mu}{\sigma}\right).$$

D'où, la densité de $Z = bY + a$, $b > 0$ et $a \in \mathbb{R}$, s'écrit en fonction de $b\mu + a$ et $b\sigma$.

En effet,

$$f_Z(z) = \frac{1}{b} f_Y\left(\frac{z - a}{b}\right) = \frac{1}{b\sigma} f_X\left(\frac{z - (b\mu + a)}{b\sigma}\right),$$

avec

$$Z = bY + a = (b\sigma)X + (b\mu + a).$$

Il est naturel de chercher une loi a priori π qui soit invariante par la transformation $g_{a,b}(\mu, \sigma) = (b\mu + a, b\sigma)$.

Soit A et B deux événements, $A \in \beta(\mathbb{R})$ et $B \in \beta(\mathbb{R}^{+*})$

$$\begin{aligned} \pi(A, B) &= P^\pi((\mu, \sigma) \in A \times B) \\ &= P^\pi(g_{a,b}^{-1}(b\mu + a, b\sigma) \in A \times B) \\ &= P^\pi((b\mu + a, b\sigma) \in g_{a,b}(A \times B)). \end{aligned}$$

Comme π doit être invariante par transformation affine alors,

$$\begin{aligned} \pi(A, B) &= P^\pi((b\mu + a, b\sigma) \in (A \times B)) \\ &= P^\pi\left((\mu, \sigma) \in \left(\frac{A-a}{b} \times \frac{B}{b}\right)\right). \end{aligned}$$

D'où,

$$\pi(A, B) = \pi\left(\frac{A-a}{b}, \frac{B}{b}\right),$$

où $\frac{A-a}{b} = \left\{\frac{\mu-a}{b}, \mu \in A\right\}$ et $\frac{B}{b} = \left\{\frac{\sigma}{b}, \sigma \in B\right\}$.

En termes de densité,

$$\begin{aligned} \int_{A \times B} \pi(\mu, \sigma) d\mu d\sigma &= \int_{\frac{A-a}{b} \times \frac{B}{b}} \pi(\mu, \sigma) d\mu d\sigma \\ &= \int_{A \times B} \pi(b\mu + a, b\sigma) b^2 d\mu d\sigma, \forall A \in \beta(\mathbb{R}) \text{ et } B \in \beta(\mathbb{R}^{+*}). \end{aligned}$$

D'où,

$$\pi(\mu, \sigma) = \pi(b\mu + a, b\sigma) b^2.$$

pour $\mu = 0$, $\sigma = \frac{1}{b}$ et $a = 0$, on obtient

$$\pi\left(0, \frac{1}{b}\right) = b^2 \pi(0, 1) \Rightarrow \pi(0, 1) = \frac{1}{b^2} \pi\left(0, \frac{1}{b}\right).$$

$\pi(0, 1)$ étant une constante, soit $\pi(0, 1) = 1$, alors,

$$\pi(0, \sigma) = \sigma^{-2}.$$

Est une loi a priori non informative impropre car

$$\begin{aligned} \int_0^{+\infty} \pi(0, \sigma) d\sigma &= \int_0^{+\infty} \sigma^{-2} d\sigma \\ &= -\sigma^{-1} \Big|_0^{+\infty} = +\infty \end{aligned}$$

2.1.4 Loi a priori de référence

La théorie a priori de référence est introduite par Bernardo (1979), est un mode général de spécification d'une loi a priori contenant aussi peu d'information que possible. Cette règle d'assignation repose sur la notion de quantité d'information d'un modèle bayésien [43].

C'est un procédé qui consiste à assigner comme une loi de probabilité a priori une densité $\pi(\theta)$ qui maximise l'information apportée par les données sur le paramètre θ , cette information est définie par:

$$I_{\underline{x}}^{\theta}(\pi) = E_{\underline{x}}(D(\pi(\theta|\underline{x}), \pi(\theta))) = \int D(\pi(\theta|\underline{x}), \pi(\theta))m(\underline{x})dx, \quad (2.1.9)$$

où $m(\underline{x})$ est la loi marginale de \underline{x} .

On mesure l'information apportée par un modèle statistique en fonction de la dépendance entre \underline{x} et θ , avec la divergence (distance) de Kullback-Leibler définie par [18]:

$$D(\pi(\theta|\underline{x}), \pi(\theta)) = \int_{\Theta} \pi(\theta|\underline{x}) \log \frac{\pi(\theta|\underline{x})}{\pi(\theta)} d\theta.$$

L'a priori $\pi(\theta)$ qui maximise la quantité $I_{\underline{x}}^{\theta}(\pi)$ vérifie l'équation:

$$\pi(\theta) \propto \exp \{L(\theta, \underline{x}) \log \pi(\theta|\underline{x})\}. \quad (2.1.10)$$

Ce problème de minimisation n'a pas de solution générale. On peut simplement vérifier que si $\pi(\theta)$ est l'a priori de référence et si $\pi(\theta|\underline{x})$ est l'a posteriori en découlant, ces deux densités doivent vérifier l'équation (2.1.10).

2.1.5 Loi a priori coïncidante - (matching prior)

Le but est de trouver une loi a priori concernant le paramètre θ qui se rapproche le plus possible de la méthode de choix fréquentiste [44].

On cherche π tel que,

$$P_{\theta}(g(\theta) \in C_x | x) = 1 - \alpha, \quad (2.1.11)$$

où C_x est un intervalle de confiance a posteriori qui est donné sur une fonction de θ , notée $g(\theta)$, unilatéral ou bilatéral. Cet ensemble définit alors un ensemble de confiance au sens fréquentiste de couverture.

$$P_{\theta}(C_x \ni g(\theta)) = \int_{C_x} (g(\theta)) f(x|\theta) dx.$$

Dans le cadre général (unidimensionnel), Welch et Peers(1963) ont démontré que lorsque $C_x = [-\infty, k_{\alpha}(x)]$

$$P_{\theta}(\theta \leq k_{\alpha}(x)) = 1 - \alpha + o\left(n^{-\frac{1}{2}}\right),$$

et que pour la loi a priori de Jeffreys

$$P_{\theta}(\theta \leq k_{\alpha}(x)) = 1 - \alpha + o\left(n^{-1}\right), \quad (2.1.12)$$

ce qui améliore l'approximation d'un facteur $\frac{1}{2}$.

Les choses se compliquent en présence de paramètre de nuisance, c'est à dire lorsque l'inférence porte sur une composante unidimensionnelle θ_1 du paramètre θ .

Le développement d'Edgeworth de la probabilité de couverture fréquentiste est donnée par:

$$P_{\theta}(\theta_1 \leq k_n(\alpha)) = 1 - \alpha + \frac{\varphi(\Phi^{-1}(1 - \alpha))}{\sqrt{n}} \left(\frac{I'(\theta) \nabla \log \pi(\theta)}{I''(\theta)^{\frac{1}{2}}} - \nabla^t \frac{I'(\theta)}{I''(\theta)^{\frac{1}{2}}} \right) + o(n^{-1}), \quad (2.1.13)$$

dans le cas unilatéral, φ et Φ sont respectivement la densité et la fonction de répartition d'une loi normale et $I(\theta)$, $I'(\theta)$ et $I''(\theta)$ sont respectivement l'information de Fisher et ses dérivées première et seconde. Dans le cas d'une région HPD bilatéral de niveau $1 - \alpha$, $C_x^{HPD}(\alpha)$ pour $\theta \in \mathbb{R}$ le développement correspondant est

$$P_{\theta}(\theta \in C_x^{HPD}) = 1 - \alpha + n^{-1} q(\alpha) b(\pi, \theta) + o\left(n^{-\frac{3}{2}}\right),$$

où q correspond à une densité du χ^2 et

$$b(\pi, \theta) = \frac{\mu'_3 - \mu''_2}{I(\theta)^2} + 2\frac{\mu'_2(\mu_3 - \mu'_2)}{I(\theta)^3} + \frac{\pi'(\theta)\mu_3 - \mu'_2}{\pi(\theta)I(\theta)^2} - \frac{\pi''(\theta)}{\pi(\theta)I(\theta)} - \frac{\mu'_2\pi'(\theta)}{\pi(\theta)I(\theta)^2},$$

les μ_j étant définis par:

$$\mu_j = E_\theta \left[\frac{\partial^j \log f(x|\theta)}{\partial \theta^j} \right], \quad j = 2, 3.$$

La loi a priori coïncidante est alors obtenue par annulation du terme d'ordre un de ce développement, comme dans l'équation différentielle de Welch et Peers(1963):

$$[I''(\theta)]^{-\frac{1}{2}} I'(\theta) \nabla \log \pi(\theta) + \nabla^t \left\{ I'(\theta) [I''(\theta)]^{-\frac{1}{2}} \right\} = 0. \quad (2.1.14)$$

Cette équation différentielle peut ne pas avoir de solution. De plus, cette solution lorsqu'elle existe, dépend du paramètre d'intérêt correspondant à ces régions HPD et diffère le plus souvent de la loi a priori de Jeffreys, même s'il existe toujours une reparamétrisation permettant de retomber sur cette dernière.

2.2 Lois a priori informatives

Une loi a priori informative peut se construire le plus objectivement possible sur la base des données historiques ou bien en introduisant une part de subjectivité. C'est-à-dire, proposer un modèle statistique cohérent combinant l'avis d'un ou plusieurs experts et construire une loi a priori [10].

2.2.1 Loi a priori conjuguée

La loi a priori conjuguée est introduite par Raiffa et Schlaifer (1961), elle peut être considérée comme un point de départ pour l'élaboration de distribution a priori fondée sur des informations a priori limitées [43]. Cette modélisation particulière par une famille paramétrée de lois a prioris est effectivement très tentante, car elle autorise des manipulations explicites des lois a posterioris.

Définition 2.2.1 (*loi a priori conjuguée*)

Une famille F de distributions de probabilité sur Θ est dite conjuguée par une fonction de vraisemblance $L(\underline{x}, \theta)$ si, pour tout $\pi \in F$, la distribution a posteriori $\pi(\theta|\underline{x})$ appartient également à F .

Remarque 2.2.1 Dans le cas de la famille conjuguée, le praticien induit directement la forme de son estimateur dès qu'il a choisi sa loi a priori [26].

Exemple 2.2.1 Soit $\underline{x} = (x_1, \dots, x_n)$ le vecteur des observations. On suppose que: $f(\underline{x}|\theta)$ est une densité d'une loi de Bernoulli de paramètre θ et que la loi a priori est une loi Bêta. En utilisant le principe de proportionnalité, on obtient que $\pi(\theta|\underline{x})$ est une bêta. On en déduit que la loi Bêta est ici conjuguée.

Les lois a priori sont généralement associées à un type particulier de loi d'échantillonnage qui permet toujours leur obtention, ces lois constituent ce qu'on appelle des familles exponentielles [3].

Lois conjuguées des familles exponentielles

On appelle famille exponentielle à s-paramètres, toute famille de lois de distribution $\{P_\theta\}$ dont la densité a la forme suivante:

$$f(x|\theta) = \exp \left[\sum_{i=1}^s \eta_i(\theta) T_i(x) - B(\theta) \right] h(x), \quad (2.2.1)$$

où $\eta_i(\cdot)$ et $B(\cdot)$ sont des fonctions du paramètre θ et $T_i(\cdot)$ sont des statistiques.

- Loi exponentielle de paramètre θ :

$$\begin{aligned} f(x|\theta) &= \frac{1}{\theta} \exp(-x|\theta) 1_{[0,+\infty]}(x). \\ &= \exp \left\{ -\frac{1}{\theta} x - \log \theta \right\} 1_{[0,+\infty]}(x) \end{aligned}$$

ici s vaut 1, $\eta_1(\theta) = \frac{1}{\theta}$, $T_1(x) = x$, $B(\theta) = \log \theta$ et $h(x) = 1_{[0,+\infty]}(x)$.

- Loi Binomiale:

$$\begin{aligned} P(X = x|\theta) &= C_n^x \theta^x (1 - \theta)^{n-x} \\ &= C_n^x \exp\{x \log \theta + (n - x) \log(1 - \theta)\} \\ &= C_n^x \exp\{x \log[\theta/(1 - \theta)] + n \log(1 - \theta)\} \end{aligned}$$

On a $s = 1$, $\eta_1(\theta) = \log[\theta/(1 - \theta)]$, $T_1(x) = x$, $B(\theta) = n \log(1 - \theta)$ et $h(x) = C_n^x$.

Le tableau ci-dessous donne un exemple de lois a priori conjuguées naturelles pour quelques familles exponentielles.

$f(x \theta)$	$\pi(\theta)$	$\pi(\theta x)$
Normale $N(\theta, \sigma^2)$	Normale $N(\mu, \tau^2)$	Normale $\left(\left(\frac{\sigma^2\mu + \tau^2x}{\sigma^2 + \tau^2} \right), \left(\frac{\sigma^2\tau}{\sigma^2 + \tau^2} \right) \right)$
Poisson $P(\theta)$	Gamma $G(\alpha, \beta)$	Gamma $G(\alpha + x, \beta + 1)$
Exponentielle (λ)	Gamma $G(\alpha, \beta)$	Gamma $G(\alpha + k, \beta + t)$
Gamma $G(\nu, \theta)$	Gamma $G(\alpha, \beta)$	Gamma $G(\alpha + \nu, \beta + x)$
Binomiale $B(n, \theta)$	Bêta $Be(\alpha, \beta)$	Bêta $Be(\alpha + x, \beta + n - x)$
Binomiale Négative $Neg(m, \theta)$	Bêta $Be(\alpha, \beta)$	Bêta $Be(\alpha + m, \beta + x)$
Multinomial $M_k(\theta_1, \dots, \theta_k)$	Dirichlet $D(\alpha_1, \dots, \alpha_k)$	Dirichlet $D(\alpha_1 + x_1, \dots, \alpha_k + x_k)$
Normale $N(\mu, 1/\theta)$	Gamma $G(\alpha, \beta)$	Gamma $G(\alpha + \frac{1}{2}, \beta + (\mu - x)^2/2)$

Tableau.2.1- Lois a priori conjuguées naturelles pour quelques familles exponentielles usuelles.

La forme canonique

Définition 2.2.2 *Un modèle exponentiel peut être écrit sous la forme, dite canonique en reparamétrisant: $\eta_i(\theta) = \theta_i$. Dans ce cas,*

$$f(x|\theta) = \exp \left[\sum_{i=1}^s \theta_i T_i(x) - A(\theta) \right] h(x). \quad (2.2.2)$$

Où $A(\cdot)$ est une fonction du paramètre θ .

On a le résultat suivant qui donne la forme des lois naturelles conjuguées dans le cas de modèle exponentiel.

Proposition 2.2.1 *Soit $f(x|\theta)$ appartenant à une famille exponentielle. Alors une famille de lois a priori conjuguées pour $f(x|\theta)$ est donnée par:*

$$\pi(\theta|\mu, \lambda) = K(\mu, \lambda) \cdot \exp(\theta\mu - \lambda A(\theta)), \quad (2.2.3)$$

où $K(\mu, \lambda)$ est une constante de normalisation, λ et μ sont appelés des hyperparamètres.

Dans ce cas, la loi a posteriori est de la forme:

$$\pi(\theta|x) \propto \exp((\mu + x)\theta - (\lambda + 1)A(\theta)). \quad (2.2.4)$$

2.2.2 Loi a priori d'entropie maximale

Si certaines caractéristiques de la loi a priori sont connues (moments, quantiles, etc), on suppose qu'elles peuvent s'écrire comme des espérances a priori,

$$E^\pi [g_k(\theta)] = w_k, \text{ avec } k = 1, \dots, K. \quad (2.2.5)$$

où g_k est une fonction de θ , $\theta \in \mathbb{R}^n$.

L'entropie a été introduite par Shannon (1948) comme une mesure de l'incertitude en théorie de l'information et en traitement du signal. Cette notion d'entropie est introduite de la manière suivante: considérons une variable aléatoire discrète X produisant des réalisations $\underline{x} = (x_1, x_2, \dots, x_n)$ et attribuons les probabilités $\pi = (\pi_1, \pi_2, \dots, \pi_n)$ à ces réalisations pour représenter notre information partielle sur cette variable. On considère la quantité $I_i = \log \left(\frac{1}{\pi_i} \right)$ comme la quantité d'information apportée par la réalisation x_i .

Le raisonnement intuitif conduisant à cette expression est que plus un événement n'est rare, plus le gain d'information obtenu par sa réalisation n'est grand. On définit alors l'entropie d'un processus par la somme pondérée des informations individuelles de chaque réalisation. C'est la définition de l'entropie dans un cadre fini et discret donnée par[17]:

$$Ent(\pi) = - \sum_{i=1}^n \pi(\theta_i) \log(\pi(\theta_i)). \quad (2.2.6)$$

La maximisation de l'entropie permet de chercher la loi qui apporte le maximum d'information. Le principe à la base de cette méthode est donc de chercher à calculer,

$$\arg \max_{\pi} Ent(\pi),$$

sous les contraintes (2.2.5).

La distribution d'entropie maximale sous les contraintes des moments (2.2.5), est la distribution associée à la densité

$$\pi^*(\theta_i) = \frac{\exp \left\{ \sum_{k=1}^K \lambda_k g_k(\theta_i) \right\}}{\sum_{j=1}^n \exp \left[- \sum_{k=1}^K \lambda_k g_k(\theta_j) \right]}, \quad j = 1, \dots, n. \quad (2.2.7)$$

La solution de ce problème est obtenue par une technique variationnelle de multiplicateurs de Lagrange. Les λ_k , sont déterminés par le système d'équations donnée par:

$$\begin{cases} -\frac{\partial \lambda_0}{\partial \lambda_k} = E[g_k(\theta)] = w_k & k = 1, \dots, K. \\ -\frac{\partial^2 \lambda_0}{\partial \lambda_k^2} = E[g_k^2(\theta)] - w_k^2 = \text{var}[g_k(\theta)] \\ -\frac{\partial^2 \lambda_0}{\partial \lambda_k \partial \lambda_l} = E[g_k^2(\theta) g_l^2(\theta)] - w_k w_l = \text{cov}\{g_k(\theta), g_l(\theta)\}. \end{cases}$$

Où

$$\lambda_0 = \ln \left\{ \sum_{j=1}^n \exp \left[- \sum_{k=1}^K \lambda_k g_k(\theta_j) \right] \right\}.$$

On peut généraliser ce qui précède au cas de distributions continues et on définit l'entropie par:

$$\text{Ent}(\pi) = - \int_{\Theta} \pi(\theta) \ln[\pi(\theta)] d\theta, \quad \text{où } \Theta \text{ est l'espace des paramètres.} \quad (2.2.8)$$

On cherche à calculer

$$\arg \max_{\pi} \text{Ent}(\pi),$$

sous les contraintes

$$\int_{\Theta} g_k(\theta) \pi(\theta) d\theta = w_k, \quad k = 1, \dots, K. \quad (2.2.9)$$

La solution est donnée par:

$$\pi^*(\theta) = \frac{\exp \left[- \sum_{k=1}^K \lambda_k g_k(\theta) \right]}{\int_{\Theta} \exp \left[- \sum_{k=1}^K \lambda_k g_k(\theta) \right] d\theta}, \quad (2.2.10)$$

les λ_k sont déterminés par le système d'équations suivant:

$$\frac{-\partial \ln \left\{ \int_{\Theta} \exp \left[- \sum_{k=1}^K \lambda_k g_k(\theta) \right] d\theta \right\}}{\partial \lambda_k} = w_k, \quad k = 1, \dots, K.$$

Le tableau suivant donne quelques exemples de lois a priori à maximum d'entropie.

Contraintes	Domaine de θ	Loi a priori $\pi(\theta)$.
$g_1(\theta) = \theta$.	\mathbb{R}_+	$\mu \exp[-\mu\theta]$ (Loi exponentielle)
$g_1(\theta) = \theta $.	\mathbb{R}	$\frac{\mu}{2} \exp[-\mu \theta]$ (Loi de Laplace)
$g_1(\theta) = \theta, g_2(\theta) = \theta^2$.	\mathbb{R}	$\frac{1}{\sqrt{2\pi\sigma^2}} \exp\left[\frac{-(\theta-m)^2}{2\sigma^2}\right]$ (Loi gaussienne)
$g_1(\theta) = \theta, g_2(\theta) = \ln \theta$.	\mathbb{R}_+	$\propto \theta^{-\lambda} \exp[-\mu\theta]$ (loi Gamma)
$g_1(\theta) = \ln \theta, g_2(\theta) = \theta^2$.	$]0, +\infty[$	$\propto \theta^{-\lambda} \exp[-\mu\theta]$ (loi de Rayleight)
$g_1(\theta) = \ln \theta, g_2(\theta) = \ln(1-\theta)$.	$]0, 1[$	$\propto \theta^{-\lambda} (1-\theta)^{-\mu}$ (loi bêta)

Tableau.2.2- Exemples de lois à maximum d'entropie.

2.2.3 Approche hiérarchique

Pour des raisons liées à la modélisation des observations ou à la décomposition de l'information a priori, il peut arriver que le modèle statistique bayésien soit hiérarchique. C'est-à-dire, mettre en jeu plusieurs niveaux de distributions a priori conditionnelles [26].

Définition 2.2.3 *Un modèle bayésien hiérarchique est un modèle statistique bayésien $(f(x|\theta), \pi(\theta))$ dans lequel la loi a priori $\pi(\theta)$ est décomposée en plusieurs lois conditionnelles, $\pi_1(\theta|\theta_1), \pi_2(\theta_1|\theta_2), \dots, \pi_n(\theta_{n-1}|\theta_n)$ et $\pi_{n+1}(\theta_n)$. Les paramètres $\theta_i, i = 1, \dots, k$, sont appelés des hyperparamètres.*

La loi a priori s'écrit,

$$\pi(\theta) = \int_{\Theta_1 \times \dots \times \Theta_n} \pi_1(\theta|\theta_1)\pi_2(\theta_1|\theta_2)\dots\pi_n(\theta_{n-1}|\theta_n)\pi_{n+1}(\theta_n)d\theta_1\dots d\theta_{n+1}, \theta_i \in \Theta_i, i = 1, \dots, n. \quad (2.2.11)$$

Considérons le cas d'un modèle hiérarchique simple où

$$\begin{cases} X \sim f(x|\theta) \\ \theta \sim \pi(\theta|\theta_1) \text{ et } \pi(\theta_1), \end{cases} \quad \text{avec } \theta \in \Theta \text{ et } \theta_1 \in \Theta_1$$

La loi a priori est donnée par:

$$\pi(\theta) = \int_{\Theta_1} \pi(\theta|\theta_1)\pi_1(\theta_1)d\theta_1. \quad (2.2.12)$$

La loi marginale de x s'écrit:

$$m(x) = \int_{\Theta} f(x|\theta)\pi(\theta)d\theta = \int_{\Theta} \int_{\Theta_1} f(x|\theta)\pi_1(\theta|\theta_1)\pi(\theta_1)d\theta_1d\theta. \quad (2.2.13)$$

D'où, la loi a posteriori de θ ,

$$\pi(\theta|x) = \frac{\int_{\Theta_1} f(x|\theta)\pi(\theta|\theta_1)\pi_1(\theta_1)d\theta_1}{\int_{\Theta} \int_{\Theta_1} f(x|\theta)\pi(\theta|\theta_1)\pi_1(\theta_1)d\theta_1d\theta}. \quad (2.2.14)$$

Remarques 2.2.2 1) Dans la plupart des cas, Θ est multidimensionnel ce qui introduit une grande complexité pour le calcul.

2) On n'a pas, en général, d'informations a priori sur l'hyperparamètre θ_1 on adopte souvent pour π_2 une distribution non informative.

3) L'intérêt essentiel de la hiérarchisation est de lever certaines difficultés calculatoires pour obtenir la loi a posteriori. En effet, dans certaines situations la loi a priori conduit

à des lois a posteriori difficile à manipuler. Utiliser la décomposition en loi conditionnelle peut permettre de lever ces difficultés.

Exemple 2.2.2 Soit $X|\theta \sim f(x|\theta)$ et $\theta \sim \exp(\lambda)$, où λ est inconnu.

La démarche hiérarchique nous conduit à considérer une loi a priori sur λ , on peut prendre par exemple une loi exponentielle de paramètre μ , avec μ est connu.

On a donc les lois suivantes : $\pi_1(\theta|\lambda) = \lambda e^{-\lambda\theta}$, $\lambda > 0$, $\theta > 0$ et $\pi_2(\lambda) = \mu e^{-\mu\lambda}$, $\lambda > 0$, $\mu > 0$.

$$\begin{aligned}\pi(\theta) &= \int_{\xi} \pi_1(\theta|\lambda)\pi_2(\lambda)d\lambda, \text{ où } \xi \text{ est l'ensemble de toutes les valeurs de } \lambda. \\ &= \int_0^{+\infty} [(\lambda \exp(-\lambda\theta)) (\mu \exp(-\mu\lambda))] d\lambda \\ &= \mu \int_0^{+\infty} \lambda [\exp(-\lambda(\theta + \mu))] d\lambda.\end{aligned}$$

Par intégration par partie, on obtient:

$$\pi(\theta) = \frac{\mu}{(\mu + \theta)^2}, \theta > 0.$$

2.2.4 Loi subjective

Il est clair que si nous ne disposons pas de données antérieures, nous devons déterminer en partie la distribution a priori $\pi(\theta)$ à l'aide d'information subjective qui est propre à l'expert, résultante d'expériences professionnelles et d'intuition raisonnable.

Par exemple, dans un cadre paramétrique cela revient à présenter des valeurs ponctuelle de θ à l'expert et pour chacune d'entre elles, de lui demander les chances qu'il lui accorde.

L'expert compare les vraisemblances intuitives des éléments de Θ de la manière suivante [49]:

$$\theta_i \text{ a } k \text{ fois plus de chances d'être réalisée que } \theta_j \implies \pi(\theta_i) = k\pi(\theta_j) \quad (2.2.15)$$

Exemple 2.2.3 Considérons un lot de n pièces et considérons la v.a. X_i définie par [37]:

Soit le modèle suivant

X_i = "le nombre de pièces défectueuses dans un lot issu de la machine numéro i , $i = 1, \dots, 5$."

Soit $p_i = \text{prob} \{ \text{Avoir une pièce défectueuse} \} = \text{proportion de pièces défectueuses}$.

alors, $X_i \sim \beta(n, p_i)$, $i = 1, \dots, 5$.

L'information a priori sur p_i est donnée par le tableau suivant:

Machine	1	2	3	4	5
Moyenne	0.3	0.4	0.5	0.2	0.2
IC.95%	[0.1; 0.5]	[0.2; 0.6]	[0.3; 0.7]	[0.05; 0.4]	[0.05; 0.4]

Tableau.2.3-Information a priori sur le paramètre p_i pour différentes machines.

Si p_i suit une loi Bêta, on ajuste les paramètres pour que la moyenne et les quartiles coïncident avec nos informations. On obtient:

Machine	1	2	3	4	5
Dist a priori	Bêta (6, 14)	Bêta (8, 12)	Bêta (12, 12)	Bêta (3.5, 14)	Bêta (3.5, 14)

Tableau.2.4- Modèle a priori sur le paramètre p_i correspondant à l'information a priori.

D'ou, $p_i \sim \text{Bêta}(\alpha_i, \beta_i)$, $i = 1, \dots, 5$.

2.3 Le poids de l'a priori dans la réponse bayésienne

Examinons cette question sur un exemple pour comprendre comment l'information a priori et l'information contenue dans les observations se combinent l'une à l'autre pour produire la réponse bayésienne. On se donne le modèle bayésien suivant:

$X_i | \theta \rightsquigarrow \text{Bernoulli}(\theta)$, $\forall i = \overline{1, n}$ et $\theta \rightsquigarrow \text{Bêta}(a, b)$. Il est commode de reparamétriser la loi bêta à l'aide de λ et μ (comme ci-dessus)

$$\begin{cases} a = \lambda\mu \\ b = \lambda(1 - \mu) \end{cases} \Rightarrow \begin{cases} \lambda = a + b \\ \mu = \frac{a}{a+b} \end{cases}$$

Où

$$\mu = E(\theta) \text{ et } V(\theta) = \frac{\mu(1-\mu)}{1+\lambda}$$

et de travailler avec la formule suivante:

$$E[\theta|\underline{x}] = \frac{\lambda}{\lambda+n} E(\theta) + \frac{n}{\lambda+n} \bar{x}.$$

En effet:

$$f(\theta|\underline{x}) \propto \frac{1}{Be(a,b)} \theta^{a-1+\sum_{i=1}^n x_i} (1-\theta)^{b+n-1-\sum_{i=1}^n x_i},$$

$$E(\theta|\underline{x}) = \frac{n\bar{x}+a}{a+b+n} = \frac{\lambda}{\lambda+n} E(\theta) + \frac{n}{n+\lambda} \bar{x}.$$

L'estimation bayésienne de θ apparaît donc comme la moyenne pondérée de \bar{x} (MV de θ), et de la moyenne a priori $E(\theta)$. Le poids de \bar{x} est proportionnel à la taille n de l'échantillon et celui de $E(\theta)$ est proportionnel à λ qui s'interprète comme la précision de l'a priori.

Géométriquement, $E(\theta|\underline{x})$ est le barycentre des points de coordonnées $E(\theta)$ et \bar{x} , affectés respectivement des coefficients $\frac{\lambda}{\lambda+n}$ et $\frac{n}{\lambda+n}$.

Cas particulier

- Si $\lambda = n$, l'estimation bayésienne de θ se situe exactement au milieu de l'intervalle $[E(\theta), \bar{x}]$.
- Si $\lambda > n$, cette estimation est plus proche de $E(\theta)$ que de \bar{x} .
- Si $\lambda < n$, cette estimation est plus proche de \bar{x} que de $E(\theta)$.

Cas limite

Pour examiner l'influence de l'a priori sur $E[\theta|\underline{x}]$ on s'intéresse aux cas limites : soit $\mu = E(\theta)$,

- pour $\lambda \rightarrow 0$ Dans le poids de l'a priori est nul, et $E[\theta|\underline{x}] \rightarrow \bar{x}$, (la réponse bayésienne tend vers l'estimation classique).
- pour $\lambda \rightarrow \infty$, le poids des données est nul et $E[\theta|\underline{x}] \rightarrow E(\theta)$ qui ne dépend plus de x (situation extrêmement informative).

Il est également intéressant de regarder ce que devient $E[\theta|\underline{x}]$ quand $n \rightarrow +\infty$, λ et μ étant fixés. Dans ce cas, le poids de l'a priori devient négligeable et la réponse bayésienne coïncide avec la réponse classique, c'est à dire \bar{x} [16].

Conclusion

Bien souvent, les connaissances subjectives sont relativement vagues et il est difficile de spécifier une loi statistique a priori précise pour les représenter. On dispose, dans ce cas, d'une certaine latitude dans le choix de cette loi a priori, mais on ne peut toutefois pas prendre n'importe quoi. Les propriétés que l'on souhaite associer à cette loi sont les suivantes :

- le calcul de la densité a posteriori à partir de la distribution a priori et la distribution de l'échantillon d'observations doit être aussi simple que possible ;
- la distribution a posteriori doit être du même type que la distribution a priori, afin de permettre un calcul d'actualisation itératif ;
- la distribution a priori doit pouvoir représenter un grand nombre de situations ;
- elle doit être possible à paramétrer et les paramètres doivent pouvoir être interprétables physiquement ;
- les règles de cohérence et de bon sens doivent être respectées.

Dans la mesure du possible, lorsque les paramètres recherchés sont constants, on choisira préférentiellement [46]:

- une densité a priori uniforme entre deux valeurs extrêmes du paramètre θ_1 et θ_2 pour θ , ou entre 0 et 1, si les connaissances initiales sont faibles, l'a priori uniforme ne favorise aucune valeur potentielle du paramètre et donne donc le maximum de poids aux observations du retour d'expérience, par contre, elle varie si l'on fait un changement de variable, ce qui est regrettable ;
- pour cela, on préfère en général choisir une distribution non informative de Jeffrey, conjuguée à la loi du phénomène observé, lorsque les connaissances a priori sont très faibles et qu'il faut favoriser les données du retour d'expérience plutôt que les connaissances a priori ;
- ou des distributions a priori conjuguées à la distribution de la fonction de vraisemblance, lorsque les connaissances initiales seront beaucoup plus informatives et que l'on souhaite une pondération entre toutes les sources d'informations disponibles.

Et lorsque les paramètres recherchés sont variables, ils seront modélisés par des densités appropriées: lois de Weibull, Gamma, Bêta, dont les paramètres de forme et d'échelle a priori suivront des lois Normale, lognormale ou uniforme, en fonctions des connaissances initiales disponibles sur ces paramètres.

Estimation Bayésienne des paramètres de la fiabilité

La fiabilité des équipements est un sujet qui préoccupe plusieurs types d'entreprises. Ces dernières ont aperçu que la fiabilité est un facteur important de leurs compétitivités. Les concepteurs et les utilisateurs des systèmes complexes montrent un grand intérêt pour les évaluations des systèmes.

Afin de caractériser la fiabilité des équipements, le fiabiliste est amené à faire des études profondes, lui permettant de déterminer les principales grandeurs de fiabilité, telles que le temps moyen de bon fonctionnement d'un matériel (ou système), son taux de défaillance, la probabilité de succès ou d'échec de sa mise en service, son temps moyen de réparation et sa disponibilité.

Pour l'estimation des paramètres de fiabilité nous utilisons les méthodes statistiques d'estimation, données par la méthode du maximum de vraisemblance et la méthode d'estimation bayésienne.

Soit T une *v.a* représentant la durée de vie d'un équipement, distribuée selon une densité f dépendant d'un paramètre inconnu θ . L'objectif est d'effectuer une inférence sur le paramètre θ .

Les analyses de sûreté, de fiabilité ou de disponibilité d'une installation s'appuient en général sur des données tirées du retour d'expérience (REX) qui sont formalisées sous

forme de banque de données élaborées dans différentes branches de l'industrie, afin de collecter l'historique du fonctionnement et des défaillances des installations.

Les banques de données de paramètres de fiabilité : Ces banques sont beaucoup plus coûteuses puisqu'elles nécessitent le suivi d'un matériel ou d'un système particulier, qui sera défini précisément dans un fichier d'identification caractérisant en particulier ses frontières, puis de collecter les informations propres au fonctionnement de ces matériels. Parmi les banques les plus connues on a la banque Européenne de donnée de fiabilité EIRIDA [46].

Dans ce présent chapitre nous appliquons la méthode bayésienne pour l'estimation des paramètres de fiabilité donnés par : la probabilité de défaillance à la sollicitation, le taux de défaillance, le temps moyen de bon fonctionnement, ainsi que les intervalles de crédibilités pour les paramètres avec utilisation de différentes lois a priori et nous comparons les résultats obtenus. À la fin de ce chapitre, nous présentons une application industrielle de données du stage effectué à l'entreprise IFRI de BEJAIA.

3.1 Les sources d'informations en fiabilité

L'estimation des caractéristiques du modèle fiabiliste d'un matériel peut être faite par des tests sur un échantillon de composants ou du matériel lui même, ou grâce au suivi du REX sur le comportement en exploitation de ces entités, et à l'aide des données issus de l'expert.

Les données disponibles sont de deux types [1]:

1. Données objectives

Ce sont des données opérationnelles issues d'origines diverses:

a. Essais sur le matériel : Il existe deux types de tests ou essais fondamentaux:

- **Les essais de type I :** La règle d'arrêt du test est un temps t prédéterminé à l'avance. Le nombre de défaillances K que l'on observera est alors une variable aléatoire ;

- **Les essais du type II :** La règle d'arrêt du test est l'observation d'un nombre prédéterminé de défaillance K , le temps d'arrêt du test T est alors une variable aléatoire.

b. Retour d'expérience : L'estimation du modèle de durée de vie à partir d'un REX observé en exploitation industrielle, vise une meilleure connaissance du comportement des

matériels de leurs modes de dégradation, de dysfonctionnement ou d'endommagement. Le REX est basé sur la collecte et la gestion des faits techniques, observé pendant toute la durée de vie du matériel, de sa mise en service jusqu'à sa désintégration.

Dans ce cas on peut avoir quatre types d'informations:

- **Données complètes** : On connaît dans ce cas l'instant t de défaillance de chaque panne ;

- **Données censurées à droite** : On décide d'arrêter l'observation à la date t_d , à cette date le composant n'a pas eu de défaillance, la date t_d sera une donnée censurée à droite, on a l'information $t_d < t$.

- **Données censurées à gauche** : On décide d'observer l'état du composant à partir de la date t_g on constate que le composant a été défaillant mais on ne sait pas à quelle date t a eu lieu cette défaillance, on a l'information $t < t_g$.

- **Données censurées par intervalle** : Le composant a eu une défaillance entre deux dates t_d et t_g connues. Il s'agit de données censurées par intervalle, on a seulement l'information $t_d < t < t_g$.

L'évaluation des paramètres de fiabilité d'un matériel à partir d'un REX est plus représentative de la réalité du terrain.

2. Données subjectives

Nous avons déjà signalé au chapitre 2, que ces informations reflètent le degré de croyance de chaque individu. Il faut noter que des données objectives sur un matériel similaire deviennent des données subjectives pour le matériel étudié.

3.2 Description du système

Considérons un système de N matériels composés de n éléments identiques non réparables. On effectue n essais indépendants sur les n éléments sur un temps cumulé t_c .

On observe un échantillon de k défaillances aux temps de défaillance t_1, t_2, \dots, t_k , et il en reste alors $(n - k)$ matériels en survies, soient aux $t_{k+1}, t_{k+2}, \dots, t_n$. D'où,

$$t_c = \sum_{i=1}^n t_i = \sum_{i=1}^k t_i + \sum_{i=k+1}^n t_j.$$

3.3 Modélisation du nombre de défaillances

Soit p la probabilité de défaillance d'un élément du système décrits ci-dessus. Pour n essais indépendants avec k occurrences de l'événement "défaillance" de probabilité p , on obtient un échantillon de REX de type binomiale [39]. Soit K une *v.a* représentant le nombre de défaillances alors K suit une loi Binomiale de paramètres n et p . D'où,

$$P(K = k) = C_n^k p^k (1 - p)^{n-k}, \quad k = \overline{1, n} \quad \text{et} \quad 0 < p < 1,$$

où,

$$p = p(T < t) = 1 - R(t), \quad t > 0.$$

La fonction de vraisemblance correspondant à l'observation de k défaillances, au cours de n essais est alors,

$$L(p, k) = P(K = k) \propto p^k (1 - p)^{n-k}. \quad (3.3.1)$$

On s'intéresse alors à estimer le paramètre p .

Estimation fréquentiste

L'estimateur classique de la probabilité de défaillance à l'instant t notée \hat{p}_{MV} , obtenu par la méthode du maximum de vraisemblance est:

$$\hat{p}_{MV} = \arg \max_p L(p, k) = \arg \max_p [\log L(p, k)].$$

On a:

$$\log L(p, k) \propto k \log p + (n - k) \log (1 - p).$$

D'où,

$$\frac{\partial}{\partial p} L(p, k) = 0 \Rightarrow \frac{k}{p} + \frac{k - n}{1 - p} = 0$$

Dont la solution est donnée par:

$$p = \frac{k}{n}.$$

Or,

$$\frac{\partial^2}{\partial p^2} L(p, k) = \frac{\partial}{\partial p} \left[\frac{k}{p} + \frac{k - n}{1 - p} \right] = - \left[\frac{k}{p^2} + \frac{n - k}{(1 - p)^2} \right] < 0.$$

D'où,

$$\hat{p}_{MV} = \frac{K}{n}. \quad (3.3.2)$$

La probabilité de survie est alors,

$$\hat{R}(t) = \hat{P}(T > t) = 1 - \hat{p}_{MV} = 1 - \frac{k}{n}. \quad (3.3.3)$$

Estimation Bayésienne: Cas non Informatif

Lorsque, on dispose d'un nombre limité d'informations sur la probabilité de défaillance p et on souhaite donner plus d'importance aux données d'essais, on choisit la loi uniforme sur $[0,1]$.

Considérons que p suit une loi a priori non informative uniforme sur $[0, 1]$, défini par:

$$\pi(p) = 1_{(0,1)}(p) = 1 \text{ pour } 0 \leq p \leq 1.$$

On obtient la loi a posteriori de p ,

$$\begin{aligned} \pi(p|k) &\propto L(p, k) \pi(p) \\ &\propto p^k (1-p)^{n-k}, \\ &\Rightarrow p|k \sim \text{Bêta}(k+1, n-k+1), \end{aligned}$$

où k est le nombre de défaillances obtenus sur les n essais.

- Le *MMSE* de la probabilité de défaillance p est:

$$\hat{p}_{MMSE} = E(p|k) = \frac{K+1}{n+2}. \quad (3.3.4)$$

- De plus, la variance a posteriori est donnée par:

$$\text{var}(p|k) = \frac{(K+1)(n-K+1)}{(n+3)(n+2)^2}.$$

L'intervalle de crédibilité de la probabilité de défaillance à la sollicitation, au niveau $(1-\gamma)$ est calculé à partir de la loi de Fisher par les expressions suivantes[39]:

$$\begin{aligned} p_{\min} &= \frac{k+1}{k+1 + (n-k+1)F_{1-\frac{\gamma}{2}}(2(n-k+1), 2(k+1))} \\ p_{\max} &= \frac{(k+1)F_{1-\frac{\gamma}{2}}(2(n-k+1), 2(k+1))}{(n-k+1) + (k+1)F_{1-\frac{\gamma}{2}}(2(n-k+1), 2(k+1))}. \end{aligned}$$

Estimation Bayésienne: Cas informatif

Supposons maintenant que nous soyons dans les conditions d'une connaissance initiale précise et informative. Par exemple, on dispose d'un REX passé d'informations provenant de banques de données ou d'essais d'experts possédant une grande expérience sur le sujet. La connaissance a priori peut être modélisée par une densité propre, conjuguée à la loi du REX. Ainsi, pour un REX binomiale on peut utiliser une loi bêta.

- **Modélisation de la loi a priori**

On considère une loi a priori pour p de type bêta à deux paramètres α et β , tels que α est le paramètre de forme (correspondant au nombre de défaillances) et β le paramètre d'échelle (correspondant au nombre d'essais). On a:

$$\pi(p) = \frac{\Gamma(\alpha + \beta)}{\Gamma(\alpha)\Gamma(\beta)} p^{\alpha-1} (1-p)^{\beta-1}, \quad 0 < p < 1, \quad \alpha, \beta > 0.$$

Avec,

$$E(p) = \frac{\alpha}{\alpha + \beta} \quad \text{et} \quad \text{var}(p) = \frac{\alpha\beta}{(\alpha + \beta)^2 (\alpha + \beta + 1)}.$$

- **Le choix des paramètres de la loi a priori Bêta**

Le choix des paramètres de la loi bêta est important et il doit être compatible avec toute l'information disponible. En général, ces paramètres peuvent être appréciés par:

- les connaissances d'experts qui proposent des valeurs raisonnables pour les valeurs extrêmes de la distribution a priori ;

- par utilisation de l'expérience passée ou par la connaissance d'une expérience similaire faite sur des matériels identiques. La forme de la distribution de la probabilité de défaillance a priori est alors connue et les paramètres seront déterminés pour présenter au mieux cette connaissance (jeunesse, maturité et usure).

- **Identification des paramètres**

On utilise la méthode des moments qui consiste à égaliser les moments théoriques a prioris avec les moments empiriques, et de résoudre le système d'équations correspondant.

Soit $\hat{\mu}_0$ le moment empirique de p , i.e,

$$\hat{\mu}_0 = \frac{1}{N} \sum_{i=1}^N p_i,$$

où N est le nombre de matériels, et p_i est la proportion empirique de panne sur le $i^{\text{ème}}$ matériel, $i = 1, \dots, n$.

Et soit $\hat{\sigma}_0^2$ la variance empirique définie par,

$$\hat{\sigma}_0^2 = \frac{1}{N} \sum_{i=1}^N (p_i - \hat{\mu}_0)^2.$$

Il s'agit de résoudre le système d'équations suivant:

$$\begin{cases} \hat{\mu}_0 = \frac{\alpha}{\alpha + \beta} \\ \hat{\sigma}_0^2 = \frac{\alpha\beta}{(\alpha + \beta)^2(\alpha + \beta + 1)}. \end{cases}$$

dont la solution est donnée par:

$$\begin{cases} \alpha = \frac{\hat{\mu}_0^2}{\hat{\sigma}_0^2} (1 - \hat{\mu}_0) - \hat{\mu}_0 \\ \beta = \frac{\hat{\mu}_0}{\hat{\sigma}_0^2} (1 - \hat{\mu}_0)^2 + \hat{\mu}_0 - 1. \end{cases} \quad (3.3.5)$$

• La loi a posteriori

La densité de la loi a posteriori de p est, par application de la règle de conjugaison, donnée par:

$$\begin{aligned} \pi(p|k) &\propto L(p, k) \pi(p) \\ &\propto p^{(\alpha+k)-1} (1-p)^{(\beta+n-k)-1}, \end{aligned}$$

où α et β sont donnés par la relation (3.3.5).

D'où,

$$p|k \sim \text{Bêta}(\alpha + k, \beta + n - k).$$

- Le *MMSE* de la probabilité de défaillance p est alors,

$$\hat{p}_{MMSE} = E(p|k) = \frac{\alpha + K}{\alpha + \beta + n}. \quad (3.3.6)$$

- De plus, la variance a posteriori est:

$$\text{var}(p|k) = \frac{(\alpha + K)(\beta + n - K)}{(\alpha + \beta + n)^2(\alpha + \beta + n + 1)}.$$

- L'intervalle de crédibilité de la probabilité de défaillance à la sollicitation au niveau $(1 - \gamma)$ est calculé en tenant compte de la règle de conjugaison et en utilisant la loi de Fisher, donné par [39]:

$$\hat{p}_{\min} = \left[\frac{(k + \alpha)}{(k + \alpha) + (n + \beta - k) F_{1-\frac{\gamma}{2}}(2n - 2k + 2\beta, 2k + 2\alpha)} \right]$$

et $\hat{p}_{\max} = \left[\frac{(k + \alpha) F_{1-\frac{\gamma}{2}}(2n - 2k + 2\beta, 2k + 2\alpha)}{(n + \beta - k) + (k + \alpha) \left(F_{1-\frac{\gamma}{2}}(2n - 2k + 2\beta, 2k + 2\alpha) \right)} \right]$.

Application numérique: Probabilité de défaillance à la sollicitation de groupes électrogènes

Dans la banque de données de fiabilité Européenne "Component Event Data Bank (CEDB)", on relève le retour d'expérience suivant [40]: sur un échantillon de 4 groupes électrogènes, on enregistre le nombre de tests effectués et le nombre de défaillances obtenues sur les éléments de ces groupes. Les données sont résumées dans le tableau suivant:

Groupe	Nbr de tests effectués	Nbr de défaillances observées
1	800	3
2	550	3
3	250	1
4	80	1
Total	1680	8

Tableau.3.1- Données du REX sur les groupes électrogènes.

- On relève un total de $k = 8$ défaillances sur $N = 4$ matériels. L'estimation fréquentiste de la probabilité de défaillance d'un élément, notée p , est donnée par:

$$\hat{p}_{MV} = \frac{k}{n} = \frac{8}{1680} = 4.76 \times 10^{-3},$$

où n est le nombre total de tests effectués sur les éléments des N groupes.

- L'estimation bayésienne, de la probabilité de défaillance des éléments des groupes électrogènes correspond au résultat d'un essai Binomial. On utilisera une loi bêta (α, β) (conjuguée) pour modéliser les données a priori.

Pour l'évaluation des paramètres de la loi bêta, on applique la méthode des moments.

Soit $\hat{\mu}_0$ la moyenne empirique, alors,

$$\hat{\mu}_0 = \frac{1}{N} \sum_{i=1}^N p_i = \frac{1}{N} \sum_{i=1}^N \frac{k_i}{n_i} = \frac{1}{4} \left(\frac{3}{800} + \frac{3}{550} + \frac{1}{250} + \frac{1}{80} \right) = 6.4 \times 10^{-3}.$$

Et soit $\hat{\sigma}_0^2$ la variance empirique. Alors,

$$\hat{\sigma}_0^2 = \frac{1}{N} \sum_{i=1}^N (p_i - \hat{\mu}_0)^2 = 0.1272 \times 10^{-4}.$$

La solution du système

$$\begin{cases} \hat{\mu}_0 = \frac{\alpha}{\alpha+\beta} \\ \hat{\sigma}_0^2 = \frac{\alpha\beta}{(\alpha+\beta)^2(\alpha+\beta+1)}. \end{cases}$$

est donnée par:

$$\begin{cases} \alpha = \frac{(0.0064)^2}{0.1272 \times 10^{-4}} (1 - 0.0064) - 0.0064 = 3.22. \\ \beta = \frac{0.0064}{0.1272 \times 10^{-4}} (1 - 0.0064)^2 + 0.0064 - 1 = 495.72. \end{cases}$$

La loi a priori de p est alors,

$$p \sim \text{Bêta}(3.22, 495.72).$$

On obtient,

$$E(p) = 6.45 \times 10^{-3} \quad \text{et} \quad \text{var}(p) = 1.28 \times 10^{-5}.$$

- La distribution a posteriori de la probabilité de défaillance à la sollicitation est alors, une $\text{bêta}(\alpha + k, \beta + n - k) = \text{bêta}(11.22, 2167.72)$.

- Le $MMSE$ de la probabilité de défaillance à la sollicitation est donnée par:

$$\hat{p}_{MMSE} = \frac{\alpha + k}{\alpha + \beta + n} = 5.15 \times 10^{-3}.$$

- De plus, la variance a posteriori est donnée par:

$$\text{var}(p|k) = \frac{(\alpha + k)(\beta + n - k)}{(\alpha + \beta + n)^2(\alpha + \beta + n + 1)} = 2.35 \times 10^{-6}.$$

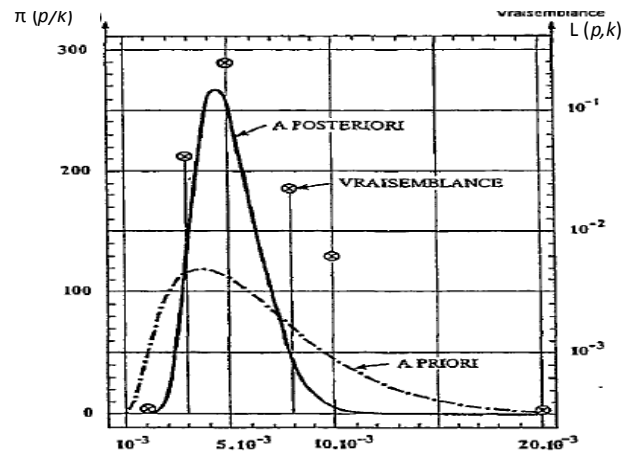


Figure.3.1- Distributions a priori et a posteriori de la probabilité de défaillance.

Interprétation de la figure: On peut constater, en particulier, que la variance a posteriori est plus faible que la variance a priori. Ce qui correspond à une diminution sensible de l'incertitude (qui est due à l'information supplémentaire apportée par le jugement d'expert) et montre que la probabilité de défaillance à la sollicitation a posteriori présente une faible dispersion.

On génère les estimations bayésiennes et fréquentielles sur les groupes électrogènes, on obtient les résultats du tableau suivant:

Groupe électrogène	Nbr de défaillance	P_{MV}	Paramètre a priori		P_{MMSE}
			α	β	
1	1	5.9×10^{-4}	0.33	1065	4.8×10^{-4}
2	0				
3	0				
4	0				
1	2	1.8×10^{-3}	0.95	879	1.5×10^{-3}
2	1				
3	0				
4	0				
1	3	4.8×10^{-3}	3.2	496	5.1×10^{-3}
2	3				
3	1				
4	1				
1	8	8.9×10^{-3}	1.28	134	9×10^{-3}
2	4				
3	2				
4	1				
1	0	4.2×10^{-4}	0.999	2347	2.5×10^{-4}
2	0				
3	0				
4	0				

Tableau.3.2- Les résultats générés pour les estimateurs fréquentistes et bayésiens.

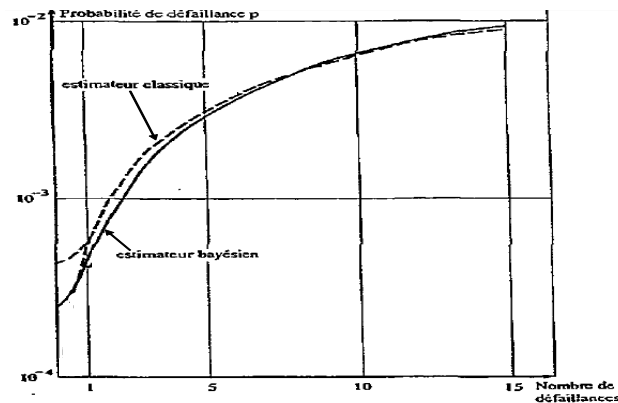


Figure.3.2- Ecart sur l'évaluation de la probabilité de défaillance entre l'inférence classique et bayésienne.

Interprétation des résultats:

- À partir du tableau et la figure ci-dessus, l'estimateur bayésien de la probabilité de défaillance des matériels est peu différent (écart 8 %) de l'estimateur fréquentiste, ce que l'on peut facilement expliquer par la quantité importante d'observations recueillies. En effet, l'écart entre l'évaluation classique et bayésienne de l'estimation de la probabilité de défaillance diminue lorsque le nombre d'observations augmente. Le résultat final est commandé essentiellement par les observations. Dans ces conditions, le choix du type de la distribution a priori n'a pas une grande importance

- Par contre, dès que le nombre d'observations diminue, l'écart relatif entre l'estimation bayésienne et l'estimation fréquentielle de la probabilité de défaillance augmente rapidement: l'estimation a priori prend un poids plus important sur le résultat final, et le choix de la distribution initiale ainsi que l'évaluation de ses paramètres d'échelle et de forme devient plus critique.

3.4 Modélisation exponentielle des temps de défaillances

La loi exponentielle joue un rôle très important pour modéliser les temps de défaillances, car elle a la particularité d'avoir un taux de défaillance constant et possède la

propriété d'absence de mémoire. Les temps de défaillances se modélisent par une loi exponentielle de paramètre λ représentant le taux de défaillance. Le temps moyen de bon fonctionnement est alors $\theta = \frac{1}{\lambda}$.

La densité de probabilité correspondante est:

$$f(t|\lambda) = \lambda \exp(-\lambda t), \quad \lambda > 0, \quad t > 0.$$

La fonction de vraisemblance de la variable λ sachant les données de l'échantillon $\underline{t} = (t_1, t_2, \dots, t_n)$ est le produit des probabilités élémentaires de l'occurrence de chacun des temps de défaillances $f(t_i|\lambda)$, $i = 1, \dots, k$, multiplié par la probabilité d'observer chacun des temps de survies constatés t_j , $j = k + 1, \dots, n$ (ce qui correspond à la fonction de fiabilité des matériels aux temps de survies) [39].

D'où,

$$\begin{aligned} L(\lambda, \underline{t}) &= \left[\prod_{i=1}^k f(t_i|\lambda) \right] \left[\prod_{j=k+1}^n R(t_j|\lambda) \right] \\ &= \prod_{i=1}^k [\lambda \exp(-\lambda t_i)] \prod_{j=k+1}^n [\exp(-\lambda t_j)] \\ &= \lambda^k \exp\left(-\lambda \sum_{i=1}^k t_i\right) \exp\left(-\lambda \sum_{i=k+1}^n t_j\right). \end{aligned}$$

On a:

$$t_c = \sum_{i=1}^k t_i + \sum_{i=k+1}^n t_j = \sum_{i=1}^n t_i.$$

D'où,

$$L(\lambda, \underline{t}) = \lambda^k \exp(-\lambda t_c) = \lambda^k \exp\left(-\lambda \sum_{i=1}^n t_i\right), \quad \lambda > 0, \quad t_i > 0, \quad i = 1, \dots, n. \quad (3.4.1)$$

Le taux de défaillance λ étant inconnu, on cherche alors à l'estimer.

Estimation fréquentiste

Pour estimer le taux de défaillance λ dans un cadre fréquentiel, on utilise la méthode du maximum de vraisemblance dont le principe est de maximiser la fonction de vraisemblance, tel que,

$$\hat{\lambda}_{MV} = \arg \max_{\lambda} L(\lambda, \underline{t}) = \arg \max_{\lambda} \log L(\lambda, \underline{t}).$$

On a:

$$\log [L(\lambda, \underline{t})] = \log \left[\lambda^k \exp \left(-\lambda \sum_{i=1}^n t_i \right) \right] = k \log \lambda - \lambda \sum_{i=1}^n t_i.$$

Et,

$$\frac{\partial}{\partial \lambda} \log [L(\lambda, \underline{t})] = 0 \Rightarrow \frac{k}{\lambda} - \sum_{i=1}^n t_i = 0.$$

D'où,

$$\lambda = \frac{k}{\sum_{i=1}^n t_i} = \frac{k}{t_c} = \frac{k}{n\bar{t}}, \text{ avec } \bar{t} = \frac{1}{n} \sum_{i=1}^n t_i.$$

$$\text{et } \frac{\partial^2}{\partial \lambda^2} \log [L(\lambda, \underline{t})] = -\frac{k}{\lambda^2} < 0.$$

- L'estimateur du M.V du taux de défaillance est alors,

$$\hat{\lambda}_{MV} = \frac{k}{n\bar{T}}, \text{ avec } \bar{T} = \frac{1}{n} \sum_{i=1}^n T_i. \quad (3.4.2)$$

- L'intervalle de confiance de l'estimateur du taux de défaillance λ au niveau $(1 - \gamma)$ est calculé à partir de la loi Khi-2. Avec,

$$\lambda_{\min} = \frac{\chi_{\gamma/2}^2(2k)}{2t_c} \text{ et } \lambda_{\max} = \frac{\chi_{1-\gamma/2}^2(2k+2)}{2t_c}.$$

- L'estimateur du M.V du temps moyen de bon fonctionnement (MTTF) est alors,

$$\hat{\theta}_{MV} = \frac{1}{\hat{\lambda}_{MV}} = \frac{\sum_{i=1}^n T_i}{k} = \frac{n\bar{T}}{k}. \quad (3.4.3)$$

Estimation Bayésienne: Cas non informatif

On considère le cas où on ne dispose pas d'informations a priori sur le paramètre à estimer λ . Alors, on choisit la loi a priori non informative de Jeffreys $\pi_J(\lambda)$ pour modéliser l'information a priori sur λ , tel que $\pi_J(\lambda) \propto [I_n(\lambda)]^{\frac{1}{2}}$, où $I_n(\lambda)$ est l'information de Fisher apportée par les observations (t_1, t_2, \dots, t_n) sur le taux de défaillance λ .

On a, à partir de (3.4.1)

$$I_n(\lambda) = -E \left[\frac{\partial^2}{\partial \lambda^2} \log [L(\lambda, \underline{t})] \right] = \frac{k}{\lambda^2}.$$

La loi a priori non informative de Jeffreys est alors,

$$\pi_J(\lambda) \propto \frac{1}{\lambda}, \quad \lambda > 0 \text{ (impropre).}$$

- La loi a posteriori du taux de défaillance λ est alors,

$$\begin{aligned} \pi(\lambda|\underline{t}) &\propto L(\lambda, \underline{t}) \pi_J(\lambda), \\ &\propto \lambda^{k-1} \exp(-\lambda t_c). \end{aligned}$$

On obtient,

$$\lambda|\underline{t} \sim \text{Gamma}(k, t_c)$$

- Le *MMSE* du taux de défaillance est la moyenne a posteriori donnée par:

$$\hat{\lambda}_{MMSE} = E(\lambda|\underline{t}) = \frac{k}{t_c} = \frac{k}{n\bar{T}} = \hat{\lambda}_{MV}. \quad (3.4.4)$$

- L'intervalle de crédibilité pour l'estimateur du taux de défaillance λ au niveau $(1 - \alpha)$ est calculé à partir de la loi Khi-2. Avec,

$$\lambda_{\min} = \frac{\chi_{\alpha/2}^2(2k)}{2t_c} \quad \text{et} \quad \lambda_{\max} = \frac{\chi_{1-\alpha/2}^2(2k)}{2t_c}.$$

- Le *MMSE* du temps moyen de bon fonctionnement est obtenu par:

$$\hat{\theta}_{MMSE} = E(\theta|\underline{t}) = \frac{t_c}{k-1} = \frac{n\bar{T}}{k-1}. \quad (3.4.5)$$

En effet,

$$\begin{aligned} E(\theta|\underline{t}) &= E(\lambda^{-1}|\underline{t}) = \int_0^{+\infty} \lambda^{-1} \pi(\lambda|\underline{t}) d\lambda \\ &= \int_0^{+\infty} \lambda^{-1} \frac{t_c^k}{\Gamma(k)} \lambda^{k-1} \exp(-\lambda t_c) d\lambda \\ &= \frac{t_c^k}{\Gamma(k)} \int_0^{+\infty} \lambda^{k-2} \exp(-\lambda t_c) d\lambda \\ &= \frac{t_c^k}{\Gamma(k)} \left[\frac{\Gamma(k-1)}{t_c^{k-1}} \right] = \frac{t_c}{k-1}. \end{aligned}$$

- Le *MMSE* de la fiabilité est donnée par la moyenne a posteriori:

$$\hat{R}_{MMSE}(t) = \left(\frac{t_c}{t_c + t} \right)^k.$$

En effet,

$$\hat{R}_{MMSE}(t) = E(R(t) | \underline{t}) = \int_0^{+\infty} \exp(-\lambda t) \pi(\lambda | \underline{t}) d\lambda$$

Estimation bayésienne: Cas informatif

- **Modélisation de l'information a priori (avis d'expert)**

Beaucoup d'experts connaissent très bien les matériels qu'ils manipulent. Ils estiment correctement la durée de vie moyenne de ces matériels grâce à leur expérience. On pourrait choisir n'importe quel type de loi présentant un maximum marqué, par exemple une loi normale centrée sur la valeur de cette durée de vie. Mais n'oublions pas que, par hypothèse, le retour d'expérience est de type exponentiel. Ces deux lois n'appartiennent pas à la même famille. Il ne semble donc pas cohérent de les fusionner. Cette fusion peut d'autre part présenter des difficultés d'intégration.

Il est raisonnable et justifié dans la démarche bayésienne de fusionner des lois d'une même famille, on considère la loi gamma (a, b) , $a > 0$ et $b > 0$, comme loi a priori conjuguée à la loi exponentielle pour le taux de défaillance λ . Elle est définie par:

$$\pi(\lambda | a, b) = \frac{b^a}{\Gamma(a)} \lambda^{a-1} \exp\{-\lambda b\}, \lambda > 0 \text{ avec } a > 0, b > 0. \quad (3.4.6)$$

Dans ces conditions, a peut être considéré comme un nombre de défaillances sur un temps cumulé de fonctionnement b .

- **Détermination des paramètres de la loi a priori gamma**

Démarche itérative

Les paramètres de la distribution a priori gamma (a, b) sont déterminés par la méthode itérative à partir des avis d'expert ($\lambda_{\min} < \lambda < \lambda_{\max}$).

Pour déterminer le nombre de défaillances a et le temps d'observation b on utilise les valeurs du taux de défaillances donnés par l'expert. On doit résoudre le système suivant[46]:

$$\begin{cases} \lambda_{\min} = \frac{\chi_{\frac{\alpha}{2}}^2(2a)}{2b} \\ \lambda_{\max} = \frac{\chi_{1-\frac{\alpha}{2}}^2(2a)}{2b}. \end{cases} \quad (3.4.7)$$

Pour résoudre ce système d'équations on doit utiliser une procédure itérative: On se donne une valeur a priori de a , avec la première équation et la valeur de λ_{\min} , on détermine b . On vérifie la cohérence de ce résultat, en reportant la valeur a priori de a et la valeur déduite de b dans la deuxième équation, et on compare le résultat avec la valeur de λ_{\max} .

Généralement, quelques itérations suffisent pour obtenir a et b .

Méthode des moments

On peut utiliser la méthode des moments pour l'estimation des paramètres de la loi gamma à partir des avis d'expert qui nous communiquent la valeur moyenne du taux de défaillance λ_{moy} et l'intervalle $[\lambda_{\min} ; \lambda_{\max}]$ [39]. Nous avons alors le système d'équation suivant:

$$\begin{cases} E[\lambda] = \lambda_{moy} = \frac{a}{b} \\ V[\lambda] = \sigma^2 = \frac{a}{b^2}. \end{cases}$$

Avec σ est tel que,

$$2 * 1.64\sigma = (\lambda_{\max} - \lambda_{\min}).$$

D'où,

$$\begin{cases} a = \frac{\lambda_{moy}^2}{\sigma^2} \\ b = \frac{\lambda_{moy}}{\sigma^2}. \end{cases} \quad (3.4.8)$$

- La loi a posteriori du taux de défaillance λ est donnée par proportionnalité:

$$\begin{aligned} \pi(\lambda|\underline{t}) &\propto L(\lambda|\underline{t})\pi(\lambda|a, b) \\ &\propto (\lambda^k \exp(-\lambda t_c)) (\lambda^{a-1} \exp(-\lambda b)). \\ &\propto \lambda^{k+a-1} \exp[-\lambda(t_c + b)]. \end{aligned}$$

D'où,

$$\lambda|\underline{t} \sim \text{Gamma}(k + a ; t_c + b). \quad (3.4.9)$$

- Le *MMSE* du taux de défaillance est la moyenne a posteriori donnée par:

$$\hat{\lambda}_{MMSE} = E(\lambda|\underline{t}) = \frac{k + a}{t_c + b}. \quad (3.4.10)$$

- L'intervalle de crédibilité symétrique bayésiens au niveau $(1 - \gamma)$ est obtenu à partir de la loi du Khi-deux, par:

$$\lambda_{\min} = \frac{\chi_{\frac{\gamma}{2}}^2(2a + 2k)}{2(b + t_c)} \text{ et } \lambda_{\max} = \frac{\chi_{(1-\frac{\gamma}{2})}^2(2a + 2k)}{2(b + t_c)}. \quad (3.4.11)$$

- Le *MMSE* du temps moyen de bon fonctionnement est donné par:

$$\hat{\theta}_{MMSE} = E(\lambda^{-1}|\underline{t}) = \int_0^{+\infty} \lambda^{-1} \pi(\lambda|\underline{t}) d\lambda.$$

Alors,

$$\hat{\theta}_{MMSE} = \frac{t_c + b}{k + a - 1}. \quad (3.4.12)$$

- Le *MMSE* de la fiabilité est donnée par la moyenne a posteriori:

$$\hat{R}_{MMSE}(t) = \left(\frac{t_c + b}{t_c + b + t} \right)^{k+a}.$$

En effet,

$$\hat{R}_{MMSE}(t) = E(R(t)|\underline{t}) = \int_0^{+\infty} \exp(-\lambda t) \pi(\lambda|\underline{t}) d\lambda$$

Application numérique: Actualisation du taux et temps de défaillance de compresseurs

Nous disposons des données de retour d'expérience (EiReDa 1988) suivant:

$k = 22$ défaillances et $t_c = 2 \times 10^5$ heures, ces observations permettent de caractériser la fonction de vraisemblance qui est par hypothèse une exponentielle [46].

- **Application pour le cas non informatif**

- Le *MMSE* du taux de défaillance est:

$$\hat{\lambda}_{MMSE} = E(\lambda|\underline{t}) = \frac{k}{t_c} = \frac{22}{2 \times 10^5} = 11 \cdot 10^{-5} \text{ défaut/heure} = \hat{\lambda}_{MV}.$$

- L'intervalle de crédibilité pour l'estimateur du taux de défaillance λ au niveau 90% est:

$$\left[\lambda_{\min} = \frac{\chi_{0.05}^2(44)}{4 \times 10^5} ; \lambda_{\max} = \frac{\chi_{0.95}^2(46)}{4 \times 10^5} \right].$$

C'est-à-dire,

$$\lambda \in [7.4 \times 10^{-5} ; 15.1 \times 10^{-5}].$$

- Le *MMSE* du temps moyen du bon fonctionnement est:

$$\hat{\theta}_{MMSE} = E(\theta|t_c, k) = \frac{t_c}{k - 1} = \frac{2 \times 10^5}{21} = 9523.81 \text{ heures.}$$

• **Application pour le cas informatif**

Les experts sont interrogés sur la valeur du taux de défaillance minimale et maximale et moyenne. Les expertises une fois agrégées et moyennées donnent: $\lambda_{\min} = 56.3 \times 10^{-6}$ défaut/heure et $\lambda_{\max} = 146 \times 10^{-6}$ défaut/heure et $\lambda_{moy} = 93 \times 10^{-6}$ défaut/heure.

Les experts estiment que ces temps sont évalués avec une bonne précision. On affecte donc un niveau de confiance de 90% à ces valeurs.

- On détermine les paramètres de la loi a priori par la méthode des moments. A partir de (3.4.7) on a:

$$\sigma = \frac{146 \times 10^{-6} - 56.3 \times 10^{-6}}{2 \times 1.64} = 27.35 \times 10^{-6}.$$

D'où,

$$b = \frac{\lambda_{moy}}{\sigma^2} = 125000 \text{ heures et } a = \frac{\lambda_{moy}^2}{\sigma^2} = 11.56 \text{ défaillances.}$$

D'où,

$$\lambda \sim \text{Gamma} (11.56 ; 125 \times 10^3).$$

Alors,

$$\lambda|t \sim \text{Gamma} (33.56 ; 325 \times 10^3).$$

- Le *MMSE* du taux de défaillance est la moyenne a posteriori donnée par:

$$\hat{\lambda}_{MMSE} = \frac{33.56}{3.25 \times 10^5} = 10.33 \times 10^{-5} \text{ défaut/heure.}$$

L'intervalle de crédibilité symétrique bayésien au niveau $(1-\gamma)$ est $[\lambda_{\inf} ; \lambda_{\sup}]$. Avec,

$$\lambda_{\min} = \frac{\chi_{0.05}^2(67, 12)}{6.5 \times 10^5} \quad \text{et} \quad \lambda_{\max} = \frac{\chi_{0.95}^2(67, 12)}{6.5 \times 10^5}.$$

D'où,

$$\lambda \in [7.58 \times 10^{-5} ; 13.42 \times 10^{-5}].$$

- Le *MMSE* du temps moyen de bon fonctionnement est donné par:

$$\hat{\theta}_{MMSE} = \frac{b + t_c}{a + k - 1} = 9622.43 \text{ heures}$$

Interprétation des résultats

	cas informatif	cas non informatif
Intervalle de crédibilité	$[7.58 \times 10^{-5} ; 13.42 \times 10^{-5}]$	$[7.44 \times 10^{-5} ; 15.12 \times 10^{-5}]$
L'erreur absolu	5.84×10^{-5}	7.68×10^{-5}

Tableau.3.3- Comparaison des intervalles de crédibilité pour le taux de défaillance.

On voit bien sur cet exemple que l'intervalle de crédibilité dans le cas bayésien informatif est plus précis que l'intervalle de crédibilité dans le cas non informatif et cela revient à l'information a priori donnée par l'expert, on peut dire que l'information a priori a amélioré l'estimation du taux de défaillance.

3.5 Modélisation Weibull des temps de défaillances

Quand on veut modéliser les temps de défaillances des matériels non réparables ou les premières pannes des matériels réparables, la loi la plus représentative dans ce cas est la loi Weibull.

Les temps de défaillances se modélisent par une Weibull à deux paramètres: η le paramètre d'échelle ou durée de vie caractéristique, c'est en quelques sorte une approximation de la durée de vie du matériel considéré, et β le paramètre de forme qui caractérise la cinétique d'évolution des défaillances. La densité de probabilité correspondante:

$$f(t) = \frac{\beta}{\eta} \left(\frac{t}{\eta}\right)^{\beta-1} \exp\left[-\left(\frac{t}{\eta}\right)^\beta\right], \quad \beta > 0, \eta > 0,$$

et la fonction de fiabilité est:

$$R(t) = \exp\left[-\left(\frac{t}{\eta}\right)^\beta\right], \quad \beta > 0, \eta > 0, t > 0.$$

La fonction de vraisemblance relative à l'observation de k défaillances survenant aux temps $t_i, i = 1, \dots, k$, et $(n - k)$ survies aux temps $t_j, j = k + 1, \dots, n$, est [19]:

$$\begin{aligned} L(\eta, \beta, \underline{t}) &= \left[\prod_{i=1}^k f(t_i|\eta, \beta) \right] \left[\prod_{j=k+1}^n R(t_j|\eta, \beta) \right], \beta > 0, \eta > 0, t > 0 \\ &= \left[\prod_{i=1}^k \frac{\beta}{\eta} \left(\frac{t_i}{\eta}\right)^{\beta-1} \exp\left[-\left(\frac{t_i}{\eta}\right)^\beta\right] \right] \left[\prod_{j=k+1}^n \exp\left[-\left(\frac{t_j}{\eta}\right)^\beta\right] \right] \end{aligned}$$

$$= \left(\frac{\beta}{\eta^\beta}\right)^k \left(\prod_{i=1}^k t_i\right)^{\beta-1} \exp \left[-\frac{1}{\eta^\beta} \left(\sum_{i=1}^k t_i^\beta + \sum_{j=k+1}^n t_j^\beta \right) \right]. \quad (3.5.1)$$

Estimation fréquentiste

Les estimateurs classiques des paramètres de la loi Weibull $\hat{\eta}_{MV}$ et $\hat{\beta}_{MV}$ sont donnés par:

$$\begin{aligned} \hat{\eta}_{MV} &= \arg \max_{\eta} L(\eta, \beta, \underline{t}) = \arg \max_{\eta} \log L(\eta, \beta, \underline{t}) \\ \text{et } \hat{\beta}_{MV} &= \arg \max_{\beta} L(\eta, \beta, \underline{t}) = \arg \max_{\beta} \log L(\eta, \beta, \underline{t}). \end{aligned}$$

On a, à partir de la relation (3.5.1),

$$\begin{aligned} \log(L(\beta, \eta, \underline{t})) &= \log \left(\left(\frac{\beta}{\eta^\beta}\right)^k \left(\prod_{i=1}^k t_i\right)^{\beta-1} \exp \left[-\frac{1}{\eta^\beta} \left(\sum_{i=1}^k t_i^\beta + \sum_{j=k+1}^n t_j^\beta \right) \right] \right) \\ &= k \log \beta + k \log \frac{1}{\eta^\beta} + (\beta - 1) \left[\sum_{i=1}^k \log t_i \right] - \frac{1}{\eta^\beta} \left[\sum_{i=1}^k t_i^\beta + \sum_{j=k+1}^n t_j^\beta \right]. \end{aligned}$$

La démonstration, faite dans (Lannoy, Procaccia, 2005) conduit à résoudre le système d'équations [19]:

$$\begin{cases} \beta = \frac{1}{\frac{\sum_{i=1}^k t_i^\beta \log t_i + (n-k)t_s^\beta \log t_s}{\sum_{i=1}^k t_i^\beta + (n-k)t_s^\beta} - \frac{1}{k} \sum_{i=1}^k t_i} \\ \eta = \left[\frac{1}{k} \left(\sum_{i=1}^k t_i^\beta + (n-k)t_s^\beta \right) \right]^{\frac{1}{\beta}}, \end{cases}$$

où t_i sont les temps de défaillance et t_s les temps de censure à droite supposés ici identiques pour tous les matériels survivants.

Ce système d'équations n'a pas de solution explicite. Elle se résout par une démarche itérative, à l'aide d'un logiciel (Rexpert, R).

• Intervalle de confiance sur les paramètres de la loi Weibull

Une des propriétés des estimateurs du M.V des paramètres de la loi Weibull est qu'ils sont asymptotiquement distribués selon une loi normale [39].

Au niveau de confiance γ , les intervalles de confiance pour les paramètres de la loi de Weibull sont donnés par :

$$\beta_{\min} = \frac{\hat{\beta}}{\exp\left(\frac{K_\gamma \sqrt{\text{var}(\hat{\beta})}}{\hat{\beta}}\right)} \quad \text{et} \quad \beta_{\max} = \hat{\beta} \exp\left(\frac{K_\gamma \sqrt{\text{var}(\hat{\beta})}}{\hat{\beta}}\right)$$

et

$$\eta_{\min} = \frac{\hat{\eta}}{\exp\left(\frac{K_\gamma \sqrt{\text{var}(\hat{\eta})}}{\hat{\eta}}\right)} \quad \text{et} \quad \eta_{\max} = \hat{\eta} \exp\left(\frac{K_\gamma \sqrt{\text{var}(\hat{\eta})}}{\hat{\eta}}\right),$$

où K_γ est le quantile d'ordre $1 - \frac{\gamma}{2}$ de la loi normale centrée réduite.

• **Intervalle de confiance sur la fiabilité**

La loi de fiabilité est égale à :

$$R(t) = \exp\left[-\left(\frac{t}{\eta}\right)^\beta\right] = \exp[-\exp(\beta(\log t - \log \eta))] = \exp[-\exp(u)],$$

avec $u = \beta(\log t - \log \eta)$.

L'intervalle de confiance au niveau $(1 - \gamma)$ de u est donné par:

$$u_{\min} = \hat{u} - K_\gamma \sqrt{\hat{u}} \quad \text{et} \quad u_{\max} = \hat{u} + K_\gamma \sqrt{\hat{u}}.$$

D'où, les bornes de l'IC au niveau $1 - \frac{\gamma}{2}$ de la fiabilité,

$$R_{\min} = \exp(-\exp(u_{\max})) \quad \text{et} \quad R_{\max} = \exp(-\exp(u_{\min})).$$

Remarque 3.5.1 Notons que le logiciel Rexpert calcul les variances de chacun des paramètres de la loi Weibull.

Estimation bayésienne du paramètre d'échelle

En pratique, il est souvent fréquent de posséder des connaissances sur le paramètre de forme β [39]. Soit parce que le phénomène physique étudié a fait l'objet de nombreux essais de vieillissement, soit parce que les experts ont une bonne connaissance de la valeur de ce paramètre par expérience. Si l'on suppose que ce paramètre est connu, on se ramène dans la configuration à un seul paramètre inconnu, comme dans le cas de la loi exponentielle.

On peut alors déterminer le paramètre d'échelle par une analogie directe avec le cas des paramètres de fiabilité constants.

On pose

$$\lambda = \frac{1}{\eta^\beta}, \quad \text{le pseudo taux de défaillance,}$$

$$\text{et } w = \sum_{i=1}^k t_i^\beta + \sum_{j=k+1}^n t_j^\beta = \sum_{i=1}^n t_i^\beta, \quad \text{le pseudo temps cumulé de défaillance.}$$

La *v.a.* T étant distribuée selon la loi Weibull de paramètres (η, β) alors la variable $w = T^\beta$ suit une distribution gamma exponentielle [39]. En effet, la fonction de vraisemblance est proportionnelle à une gamma exponentielle, donnée par :

$$L(\eta, \beta, \underline{t}) \propto \lambda^k \beta^k \left(\prod_{i=1}^k t_i \right)^{\beta-1} \exp(-\lambda w)$$

$$\propto \lambda^k \exp(-\lambda w), \quad \lambda > 0, \quad w > 0. \quad (3.5.2)$$

Cas non informatif

On suppose que les connaissances a priori sont faibles, et que l'on souhaite favoriser l'impact du retour d'expérience dans le résultat final. Alors, dans ce cas, on choisit la loi a priori de Jeffreys, définie par:

$$\pi_J(\lambda) = \frac{1}{\lambda}, \quad \lambda > 0 \quad (\text{impropre}). \quad (3.5.3)$$

- À partir de la relation (3.5.2) et (3.5.3), la loi a posteriori de λ , est alors,

$$\begin{aligned} \pi(\lambda|\underline{t}) &\propto L(\lambda, \underline{t}) \pi_J(\lambda) \\ &\propto \lambda^k \frac{1}{\lambda} \exp(-\lambda w) \\ &\propto \lambda^{k-1} \exp(-\lambda w), \quad \lambda > 0. \end{aligned}$$

D'où,

$$\lambda|\underline{t} \sim \text{Gamma}(k, w).$$

- Le *MMSE* du pseudo taux de défaillance de la loi Weibull est défini par:

$$\hat{\lambda}_{MMSE} = \frac{k}{w} = \frac{k}{\sum_{i=1}^n t_i^\beta}$$

- Les $MMSE$ du paramètre d'échelle de la loi Weibull est défini par:

$$\hat{\eta}_{MMSE} = \frac{w^{1/\beta} (\Gamma(k - 1/\beta))}{\Gamma(k)}.$$

En effet,

$$\hat{\eta}_{MMSE} = E(\eta|\underline{t}) = E\left(\lambda^{-1/\beta}|\underline{t}\right) = \int_0^{+\infty} \lambda^{-1/\beta} \pi(\lambda|\underline{t}) d\lambda.$$

Le temps moyen de bon fonctionnement $MTTF$ pour la loi Weibull étant θ égal à:

$$MTTF = \theta = \lambda^{-1/\beta} \Gamma\left(1 + \frac{1}{\beta}\right).$$

- Le $MMSE$ de θ , est alors,

$$\hat{\theta}_{MMSE} = \frac{w^{1/\beta} \Gamma(k - 1/\beta) \Gamma(1 + 1/\beta)}{\Gamma(k)}. \quad (3.5.4)$$

En effet,

$$\hat{\theta}_{MMSE} = E(\theta|\underline{t}) = \int_0^{+\infty} \theta \pi(\lambda|\underline{t}) d\lambda = \int_0^{+\infty} \lambda^{-1/\beta} \Gamma\left(1 + \frac{1}{\beta}\right) \frac{w^k}{\Gamma(k)} \lambda^{k-1} \exp(-\lambda w) d\lambda.$$

- L'intervalle de crédibilité pour θ est symétrique au niveau $(1 - \alpha)$ et peut être calculé à l'aide de la loi du Khi-2, par [39]:

$$\hat{\theta}_{\min} = \Gamma(1 + 1/\beta) \left[\frac{\chi_{(1-\alpha/2)}^2(2k)}{2w} \right]^{-1/\beta} \quad \text{et} \quad \hat{\theta}_{\max} = \Gamma(1 + 1/\beta) \left[\frac{\chi_{(\alpha/2)}^2(2k)}{2w} \right]^{-1/\beta}.$$

Cas informatif

Supposons que cette fois-ci, les experts ont des connaissances précises sur la valeur du taux de défaillance du matériel (ou sa pseudo durée de vie caractéristique γ), ce qui permet de définir une densité a priori gamma de paramètres (a_0, b_0) pour λ . Où a_0 et b_0 seront obtenus à partir de la connaissance d'une valeur minimale et maximale du taux de défaillance par les experts, en résolvant le système d'équations associé.

Les pseudos taux de défaillance λ_{\min} et λ_{\max} sont donnés par:

$$\lambda_{\min} = \frac{1}{\gamma_{\max}} = \frac{1}{\eta_{\max}^\beta} \quad \text{et} \quad \lambda_{\max} = \frac{1}{\gamma_{\min}} = \frac{1}{\eta_{\min}^\beta}.$$

où γ est la pseudo durée de vie caractéristique a priori.

Remarque 3.5.2 En général, il est préférable d'interroger les experts sur la valeur de la durée de vie moyenne d'un matériel plutôt que sur la durée de vie caractéristique.

- La loi a posteriori de λ , par application de la règle de conjugaison est aussi une Gamma de paramètres $(a_0 + k, b_0 + w)$.
- On en déduit que la distribution a posteriori de γ est une Gamma inverse de paramètres $(a_0 + k, (b_0 + w)^{-1})$, puisque $\gamma = \frac{1}{\lambda}$.
- Le *MMSE* du pseudo taux de défaillance est:

$$\hat{\lambda}_{MMSE} = \frac{a_0 + k}{b_0 + w}. \quad (3.5.5)$$

- Le *MMSE* du paramètre d'échelle η est:

$$\hat{\eta}_{MMSE} = \frac{\Gamma(a_0 + k - 1/\beta) (b_0 + w)^{1/\beta}}{\Gamma(a_0 + k)}.$$

En effet,

$$\hat{\eta}_{MMSE} = E\left(\lambda^{-1/\beta} | \underline{t}\right) = \int_0^{+\infty} \lambda^{-1/\beta} \pi(\lambda | \underline{t}) d\lambda. \quad (3.5.6)$$

- Le *MMSE* du temps moyen de bon fonctionnement est:

$$\hat{\theta}_{MMSE} = \frac{\Gamma(1 + 1/\beta) \Gamma(a_0 + k - 1/\beta) (b_0 + w)^{1/\beta}}{\Gamma(a_0 + k)}.$$

En effet,

$$\hat{\theta}_{MMSE} = E(\theta | \underline{t}) = \int_0^{+\infty} \theta \pi(\lambda | \underline{t}) d\lambda = \int_0^{+\infty} \lambda^{-1/\beta} \Gamma\left(1 + \frac{1}{\beta}\right) \pi(\lambda | \underline{t}) d\lambda. \quad (3.5.7)$$

- L'intervalle de crédibilité au niveau $(1 - \alpha)$ à partir de la loi du Khi-2 pour $\hat{\theta}_{MMSE}$ est [39]:

$$\hat{\theta}_{\min} = \Gamma(1 + 1/\beta) \left[\frac{\chi_{(1-\alpha/2)}^2 (2k + 2a_0)}{(2w + 2b_0)} \right]^{-\frac{1}{\beta}}$$

et

$$\hat{\theta}_{\max} = \Gamma(1 + 1/\beta) \left[\frac{\chi_{(\alpha/2)}^2 (2k + 2a_0)}{2w + 2b_0} \right]^{-\frac{1}{\beta}}.$$

Application numérique: Evaluation du vieillissement d'un échangeur de chaleur

Le retour d'expérience d'un échangeur de chaleur fait apparaître les temps de défaillances suivants (en années):

0.5 ; 0.6 ; 0.75 ; 0.90 ; 1.10 ; 1.15 ; 1.20 ; 1.30 ; 1.35 (d'après Martz et Waller).

Compte tenu du REX sur d'autres installations, les exploitants savent que le phénomène de dégradation dominant sur ces échangeurs est de la corrosion sous tension du faisceau tubulaire. Dans ces conditions, la loi de vieillissement de ces matériels est bien modélisée par une loi weibull ayant un paramètre de forme β élevé et estimé à 3.4 (d'après les essais effectués au laboratoire sur éprouvettes) [39].

Pour simplifier le problème, on admet que ces estimations correspondent au paramètre caractéristique de la loi Weibull et non pas à la durée de vie moyenne du faisceau tubulaire (Martz et Waller, 1992).

Modélisation du problème

On calcul successivement :

- Le pseudo temps cumulé de test $k = n = 9$ défaillances (essai complet):

$$w = \sum_{i=1}^n t_i^\beta = (0.5)^{3.4} + (0.6)^{3.4} + \dots + (1.35)^{3.4} = 11.4 \text{ ans.}$$

• Application pour le cas non informatif

- Le *MMSE* du pseudo taux de défaillance est:

$$\hat{\lambda}_{MMSE} = E(\lambda|t) = \frac{k}{w} = 0.79 \text{ défaut/année.}$$

- D'où, le *MMSE* du paramètre d'échelle η est:

$$\hat{\eta}_{MMSE} = \frac{w^{1/\beta} \Gamma(k - 1/\beta)}{\Gamma(k)} = \frac{(11.4)^{1/3.4} 21600}{40320} = 1.096 \text{ année.}$$

- Le *MMSE* du temps moyen du bon fonctionnement $\hat{\theta}_{MMSE}$ est:

$$\hat{\theta}_{MMSE} = \frac{w^{1/\beta} \Gamma(k - 1/\beta) \Gamma(1 + 1/\beta)}{\Gamma(k)} = \frac{(11.4)^{1/3.4} \times 21600 \times 0.8984}{40320} = 0.98 \text{ année.}$$

- L'intervalle de crédibilité de $\hat{\theta}_{MMSE}$ au niveau 90% et au seuil de 10% est défini à partir de la loi du Khi-2 [39]:

$$\hat{\theta}_{\min} = \Gamma(1 + 1/\beta) \left[\frac{\chi_{(1-\alpha/2)}^2(2k)}{2w} \right]^{-1/\beta} = 0.8373, \text{ avec } \chi_{(1-\alpha/2)}^2(2k) = 28.97.$$

$$\text{et } \hat{\theta}_{\max} = \Gamma(1 + 1/\beta) \left[\frac{\chi_{(\alpha/2)}^2(2k)}{2w} \right]^{-1/\beta} = 1.1662, \text{ avec } \chi_{(\alpha/2)}^2(2k) = 9.39.$$

Soit finalement:

$$0.8373 \leq \hat{\theta}_{MMSE} \leq 1.1662.$$

• **Application pour le cas informatif**

Les paramètres de la distribution a priori Gamma pour λ sont estimés par les experts, par les valeurs $a_0 = 2$ et $b_0 = 2$.

- Le *MMSE* du pseudo taux de défaillance est:

$$\hat{\lambda}_{MMSE} = \frac{a_0 + k}{b_0 + w} = 0.82 \text{ défaut/année.}$$

D'où, le *MMSE* du paramètre d'échelle est:

$$\hat{\eta}_{MMSE} = \frac{\Gamma(a_0 + k - 1/\beta) (b_0 + w)^{1/\beta}}{\Gamma(a_0 + k)} = \frac{1.825 \times 10^6 \times 2.145}{3628800} = 1.079 \text{ année.}$$

- Le *MMSE* du temps moyen de bon fonctionnement est:

$$\hat{\theta}_{MMSE} = \frac{\Gamma(1 + 1/\beta) \Gamma(a_0 + k - 1/\beta) (b_0 + w)^{1/\beta}}{\Gamma(a_0 + k)} = 0.97 \text{ année.}$$

- L'intervalle de crédibilité au niveau 90% de $\hat{\theta}_{MMSE}$, déterminé à l'aide de la loi du Khi-2, par:

$$0.8382 \leq \hat{\theta}_{MMSE} \leq 1.1286.$$

• **Interprétation des résultats**

	cas informatif	cas non informatif
Intervalle de crédibilité	[0.8382 ; 1.1286]	[0.8373 ; 1.1662]
L'erreur absolu	0.2904	0.3289

Tableau.3.4- Comparaison des intervalles de crédibilité pour θ .

Les résultats d'estimations obtenus dans le cas informatif et non informatif sont légèrement différents. L'intervalle de crédibilité dans le cas informatif est plus précis. Ce qui peut s'expliquer par l'informativité relative du REX, vis-à-vis de la connaissance a priori.

3.6 Application aux données du stage

Nous avons, procédé à une enquête d'une période d'un mois au sein de l'entreprise IFRI «IBRAHIM & FILS» qui se situe à Ighzar Amokrane, dont le domaine d'activité est la production et la commercialisation des boissons diverses.

Durant notre stage nous avons interrogé les exploitants et les concepteurs du service maintenance et méthode, dans le but de collecter les informations nécessaires afin de déterminer le champ de notre étude. De ce fait, nous avons choisi pour l'étude la machine souffleuse car c'est la machine la plus fragile et provoque des arrêts importants voire annexe B.

3.6.1 Collecte et constitution d'échantillons de données

Pour reconstituer l'échantillon des données du REX concernant les observations des temps de défaillances et les temps de survies de la machine souffleuse, nous avons cherché toutes les données nécessaires à la présente étude: les données de base concernant les dates de défaillances et les écarts de temps entre les défaillances. Une information supplémentaire sur la valeur minimale et maximale du taux de défaillance de la machine souffleuse a été obtenue. Cette information représente l'information subjective propre aux concepteurs (experts) de l'entreprise ayant une expérience sur le fonctionnement de la machine. On cherche à déterminer l'impact de cette information sur l'estimation.

3.6.2 Traitement statistique des données

Etant donné que les défaillances surviennent d'une façon aléatoire, alors t_1, t_2, \dots, t_k , sont les observations de la variable aléatoire temps de défaillance de la machine souffleuse, pour laquelle nous associerons la loi exponentielle de paramètre λ , car c'est la loi la plus représentative dans notre cas, Le test d'ajustement de Kolmogorov-Smirnov nous a donné une statistique égale à 0.126 inférieur à la valeur tabulée qui est égale à 0.15 (voir la figure ci-dessous). D'où, nous validons le modèle exponentiel.

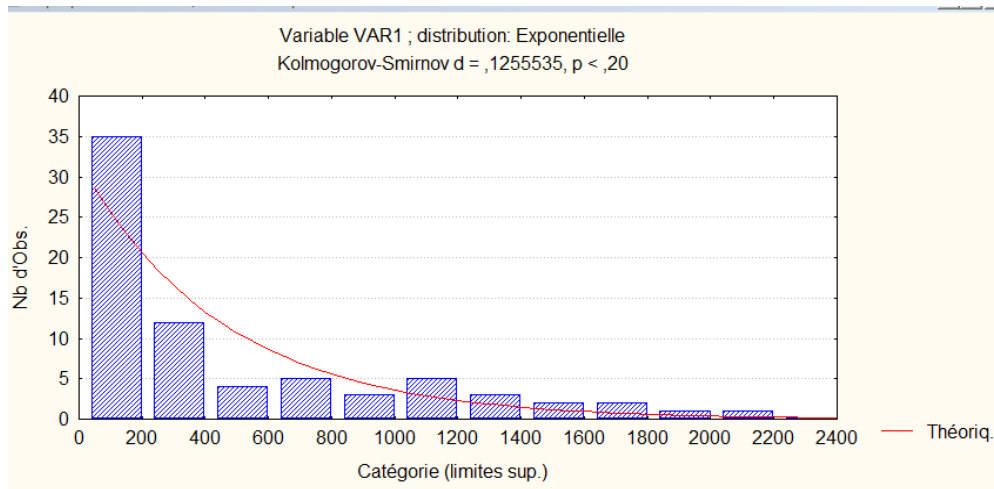


Figure.3.3- Ajustement des temps de défaillances de la machine souffleuse par la loi exponentielle.

Sur $n = 146$ observations, nous avons observé $k = 73$ défaillances aux temps (t_1, t_2, \dots, t_k) et $(n - k)$ survies au temps (t_{k+1}, \dots, t_n) . Le temps cumulé t_c est alors,

$$t_c = \sum_{i=1}^{73} t_i + \sum_{j=73+1}^{146} t_j = 1484 + 33425 = 34907mn \simeq 581.783 \text{ heure.}$$

3.6.3 Estimation des paramètres de la fiabilité

Estimation fréquentiste

- Le taux de défaillance est égale à:

$$\hat{\lambda}_{MV} = \frac{k}{t_c} = \frac{73}{581.783} = 0.1255 \text{ défaut/heure} = 3.012 \text{ défaut/jour}$$

- Le temps moyen de bon fonctionnement est égale à:

$$MTTF = \hat{\theta}_{MV} = \frac{1}{\hat{\lambda}_{MV}} = 7.97 \text{ heures.}$$

- L'intervalle de confiance pour λ au niveau de confiance $1 - \alpha = 90\%$ est :

$$\begin{cases} \lambda_{\min} = \frac{\chi_{\alpha/2}^2(2k)}{2t_c} = \frac{\chi_{0.05}^2(2 \times 73)}{2 \times 581.783} = \frac{119.05}{1163.566} = 0.1023 \text{ défaut/heure} = 2.4552 \text{ défaut/jour} \\ \lambda_{\max} = \frac{\chi_{1-\alpha/2}^2(2k+2)}{2t_c} = \frac{177.45}{1163.566} = 0.1525 \text{ défaut/heure} = 3.66 \text{ défaut/jour.} \end{cases}$$

$$2.46 \leq \hat{\lambda}_{MV} \leq 3.66.$$

- La fonction de fiabilité estimée est:

$$\hat{R}(t) = \exp\left(-\hat{\lambda}_{MV}t\right) = \exp(-0.1255 t), \quad t > 0.$$

Estimation Bayésienne

- **Modélisation de la loi a priori**

Nous avons interrogé les concepteurs du service méthode sur le taux de défaillance de la machine souffleuse, ils nous ont donné une valeur minimale $\lambda_{\min} = 2$ défaut/jour et une valeur maximale $\lambda_{\max} = 8$ défaut/jour.

Les concepteurs sont très informatifs compte tenu de leurs grandes expériences, ils estiment que les taux de défaillances sont évalués avec une bonne précision. De plus on dispose d'un REX exponentiel, alors on est amené à choisir une loi a priori informative de type gamma de paramètres a et b .

Identification des paramètres de la loi Gamma

Pour déterminer les paramètres de la loi Gamma, on utilise la démarche itérative à partir des deux valeurs du taux de défaillances λ_{\min} et λ_{\max} .

On sait qu'à partir de la loi du Khi-2 et un niveau de confiance de $1 - \alpha = 90\%$, on a un système de deux équations avec deux inconnus a et b .

$$\begin{cases} \lambda_{\min} = \frac{\chi_{\alpha/2}^2(2a)}{2b} \\ \lambda_{\max} = \frac{\chi_{1-\alpha/2}^2(2a)}{2b}. \end{cases}$$

On a effectué la procédure d'itération. On a initialisé k à 1 puis, en augmentant la valeur de k , par approximations successives, on arrive à $k = 6$. On trouve respectivement,

$$\chi_{0.05}^2(2 \times 6) = 5.23 \quad \text{et} \quad \chi_{0.95}^2(2 \times 6) = 21.03.$$

On en déduit avec la première équation que:

$$b = \frac{5.23}{2 \times 2} = 1.31 \text{ jours} = 31.44 \text{ heures.}$$

D'où,

$$\lambda_{\max} = \frac{\chi_{0.95}^2(2 \times 6)}{2 \times 1.31} = \frac{21.03}{2.62} = 8.03 \simeq 8 \text{ défaut/jour.}$$

Ce résultat étant accepté, on adopte les valeurs $a = 6$ défaillances et $b = 31.44$ heures. On obtient:

$$\lambda \sim \text{Gamma}(6 ; 31.44).$$

• **La loi a posteriori**

Par application de la règle de conjugaison, la loi a posteriori de λ est aussi une gamma de paramètres $k + a = 73 + 6 = 79$ défaillances et $b + t_c = 31.44 + 581.783 = 613.223$ heures. D'où,

$$\lambda|t \sim \text{Gamma}(79 ; 613.223).$$

- Le *MMSE* du taux de défaillance est:

$$\hat{\lambda}_{MMSE} = \frac{k + a}{b + t_c} = \frac{79}{613.223} = 0.1288 \text{ défaut/heure} = 3.09 \text{ défaut/jour.}$$

- Le *MMSE* du temps moyen de bon fonctionnement est:

$$\hat{\theta}_{MMSE} = \frac{b + t_c}{a + k - 1} = \frac{613.223}{78} = 7.86 \text{ heures.}$$

- L'intervalle de crédibilité pour λ au niveau de confiance 90% est:

$$\begin{aligned} \lambda_{\min} &= \frac{\chi_{\alpha/2}^2(2(a+k))}{2(b+t_c)} = \frac{\chi_{0.05}^2(2(6+73))}{2(31.44+581.783)} = \frac{129.91}{1226.446} \\ &= 0.1059 \text{ défaut/heure} = 2.542 \text{ défaut/jour.} \end{aligned}$$

$$\lambda_{\max} = \frac{\chi_{1-\alpha/2}^2(2(a+k))}{2(b+t_c)} = \frac{188.39}{1226.446} = 0.1536 \text{ défaut/heure} = 3.686 \text{ défaut/jour.}$$

$$2.542 \leq \hat{\lambda}_{MMSE} \leq 3.686.$$

- La fonction de fiabilité estimée est:

$$\hat{R}_{MMSE}(t) = \left(\frac{t_c + b}{t_c + b + t} \right)^{k+a} = \left(\frac{613.223}{613.223 + t} \right)^{79}.$$

• **Interprétation des résultats**

	Intervalle de confiance	Intervalle de crédibilité
	[2.46; 3.66]	[2.542 ; 3.686]
Erreure absolu	1.2	1.1

Tableau. 3.5- Comparaison des intervalles de crédibilité pour λ .

Les résultats d'estimation obtenus par les méthodes fréquentiste et bayésienne sont légèrement différents. L'intervalle de crédibilité pour λ est plus précis par rapport à l'intervalle de confiance fréquentiste. Cela est dû essentiellement à l'information a priori apportée par les jugements des concepteurs de l'entreprise. Cette information a amélioré légèrement les résultats.

Ce qu'on peut déduire sur la machine souffleuse, c'est qu'elle est en bon état de fonctionnement. C'est-à-dire, elle est dans sa vie utile, et que son temps moyen de bon fonctionnement est approximativement égale à $7.86 \simeq 8$ heures après chaque défaillance (d'après la propriété sans mémoire de la loi exponentielle). Cependant, les concepteurs de l'entreprise doivent améliorer leurs conditions de travail pour éviter les micros arrêts.

Conclusion

Dans ce chapitre, nous avons considéré trois applications et une autre pour les données du stage.

Nous avons d'abord considéré un système dans sa période de jeunesse, pour lequel nous avons associé le modèle exponentiel ensuite le modèle weibull pour modéliser les différentes phase de vie du système. Ainsi que la modélisation du nombre de défaillance où nous avons estimé la probabilité de défaillance à la sollicitation, avec la méthode classique et bayésienne.

L' application concernant les données du stage sur la machine souffleuse à laquelle nous avons associé le modèle exponentiel, ce qui permet de conclure que la machine est dans sa période de maturité.

Les estimations effectuées dans les applications nous ont permis de conclure que l'approche bayésienne nous propose une meilleure estimation, quelque soit la loi a priori considérée. L'information a priori lorsqu'elle est importante doit apporter un plus à l'estimation dans la précision.

Conclusion Générale

Ce travail nous a permis d'étudier l'impact de la statistique bayésienne dans le domaine de la fiabilité. Cette méthode est l'une des techniques les plus explorées à l'heure actuelle, dans la mesure où elle permet d'estimer les paramètres de fiabilité à partir des connaissances antérieures et des avis d'experts, représentant une information a priori qui est exprimée sous forme d'une loi de probabilité afin de générer une nouvelle information.

Dans le cadre de ce travail, nous avons expliqué comment traduire l'information a priori à une loi de probabilité, dite a priori. Nous avons considéré deux approches pour l'estimation bayésienne. À savoir, l'approche informative et l'approche non informative. Nous avons constaté que l'estimation bayésienne apporte une amélioration dans le cas informatif relativement à l'importance de l'information a priori.

Dans le cas échéant elle nous donne la même estimation que celle obtenu par l'approche classique.

Nous avons confirmé ce résultat dans les différentes applications que nous avons considéré dans le domaine de fiabilité, pour la détermination des paramètres de fiabilité à savoir: la probabilité de défaillance à la sollicitation, le taux de défaillance et le temps moyen de bon fonctionnement.

Une autre application est effectuée avec les données du stage pratique (Entreprise IFRI de BEJAIA), où nous avons considéré le cas informatif dans l'estimation bayésienne. L'information a priori donnée par les experts de l'entreprise IFRI de BEJAIA a été modélisée sous forme d'une loi gamma et nous a conduit à améliorer l'estimation fréquentiste obtenue pour le taux de défaillance de la machine souffleuse. Cette amélioration est de l'ordre de 10%.

En terme de perspectives, il est souhaitable d'envisager plusieurs voies de développement: une étude de fiabilité avec des données d'essais, modéliser les lois a prioris à travers

les données passées et d'appliquer les méthodes MCMC dans le cas des lois a priori non conjugués ou difficile à intégrer.

Annexe A

D'autres modèles de fiabilité

Lois continues

Loi Normale ou Loi de Laplace-Gauss

La loi Normale est la loi la plus répandue parmi les lois de probabilité car elle s'applique à de nombreux phénomènes, notamment en physique et en économie (erreurs de mesure).

C'est une loi absolument continue et symétrique dépendant de deux paramètres μ et σ^2 , où μ est la moyenne (qui correspond au paramètre de position), et σ^2 la variance (le paramètre d'échelle qui mesure la dispersion de la variable aléatoire autour de sa moyenne μ). Elle est notée $N(\mu, \sigma^2)$.

Cette loi s'utilise pour représenter la distribution des durées de vie de dispositifs en fin de vie (usure), car le taux de défaillance est croissant.

- **Densité de probabilité**

$$f(t) = \frac{1}{\sigma\sqrt{2\pi}} \exp \left[-\frac{1}{2} \left(\frac{t - \mu}{\sigma} \right)^2 \right], \quad t \in \mathbb{R}.$$

- **Fonction de répartition**

$$F(t) = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^t \exp \left[-\frac{1}{2} \left(\frac{x - \mu}{\sigma} \right)^2 \right] dx, \quad t \in \mathbb{R}, \quad \mu \in \mathbb{R}, \quad \text{et } \sigma > 0.$$

- **Fonction de fiabilité**

$$R(t) = 1 - F(t) = \frac{1}{\sigma\sqrt{2\pi}} \int_t^{\infty} \exp \left[-\frac{1}{2} \left(\frac{x - \mu}{\sigma} \right)^2 \right] dx, \quad t > 0.$$

- **MTTF**

$$E(T) = \int_{-\infty}^{+\infty} t f(t) dt = \mu.$$

- **La variance**

$$V(T) = E[(t - \mu)^2] = \int_{-\infty}^{+\infty} [(t - E(t))^2] f(t) dt = \sigma^2.$$

La loi Normale $N(0, 1)$ est appelée loi Normale centrée réduite et considérée comme loi de référence. La loi Normale $N(0, 1)$ est tabulée.

Si T suit une loi Normale $N(\mu, \sigma^2)$ alors $Y = \frac{T - \mu}{\sigma}$ suit une loi Normale centrée réduite $N(0, 1)$.

Loi du Khi-2

La loi de la variable aléatoire continue $T = \sum_{i=1}^{\nu} X_i^2$, sommes des carrés de ν variable X_i Normales centées réduites et indépendantes est une loi Khi-2 à ν degrés de liberté (*d.d.l*)

- **Densité de probabilité**

La densité de probabilité (qui dépend du paramètre ν) est :

$$f(t) = \frac{1}{2^{\frac{\nu}{2}} \Gamma\left(\frac{\nu}{2}\right)} (t)^{\frac{\nu}{2}-1} \exp\left(-\frac{t}{2}\right), \quad t > 0.$$

- **La moyenne**

$$E(T) = \nu.$$

- **La variance**

$$var(T) = 2\nu.$$

Cette loi est tabulée et très utilisée pour déterminer un intervalle de confiance (*IC*) d'un taux de défaillance dans la démarche fréquentielle. Comme cette loi est celle de la somme au carré de variables Normales centrées réduites, le *d.d.l* est égale à deux fois le nombre d'observations k , soit deux fois le nombre de défaillances $2k$.

Si l'on veut déterminer un IC symétrique avec un risque α , la surface de la distribution comprise entre les bornes de IC sera égale à $(1 - \alpha) \%$ de la surface totale de la distribution Normale. La borne inférieure de IC correspond donc au quantile d'ordre $(\alpha/2)$ et la borne supérieure au quantile d'ordre $(1 - \alpha/2)$ de la distribution. Le $d.d.l$ étant par ailleurs égale à $\nu = 2k$ [39].

Loi log-Normale (ou loi de Galton)

Une variable aléatoire T est distribuée suivant la loi log-normale de paramètre $\mu \in \mathbb{R}$, et σ^2 , $\sigma > 0$ et on note $T \sim LN(\mu, \sigma^2)$, si son logarithme est distribué suivant une loi Normale de paramètres μ et σ^2 . Cette distribution est largement utilisée pour modéliser des durées de vie des matériaux, en particulier le vieillissement, ou pour représenter la distribution de certaines caractéristiques des matériaux (leurs propriétés mécaniques) et les temps de réparations.

C'est un modèle approprié pour représenter par exemple les défaillances causées par des fissures dues à la fatigue. Il peut donc être utilisé pour représenter plus généralement, le temps de repture à la suite de la propagation de fissure dans un matériel[39].

- **Densité de probabilité**

$$f(t) = \frac{1}{\sigma t \sqrt{2\pi}} \exp \left[-\frac{1}{2} \left(\frac{\log t - \mu}{\sigma} \right)^2 \right], \quad t > 0, \quad \sigma > 0, \quad \mu \in \mathbb{R}.$$

- **Fonction de répartition**

$$F(t) = \frac{1}{\sigma \sqrt{2\pi}} \int_0^t \frac{1}{x} \exp \left[-\frac{1}{2} \left(\frac{\log x - \mu}{\sigma} \right)^2 \right] dx, \quad t > 0.$$

- **Fonction de fiabilité**

$$R(t) = 1 - \frac{1}{\sigma \sqrt{2\pi}} \int_0^t \frac{1}{x} \exp \left[-\frac{1}{2} \left(\frac{\log x - \mu}{\sigma} \right)^2 \right] dx = 1 - \Phi \left(\frac{\log t - \mu}{\sigma} \right), \quad t > 0,$$

où, Φ est la fonction de répartition de la loi normale centrée réduite.

- **Taux de défaillance**

$$\lambda(t) = \frac{\exp \left[\frac{-1}{2} \left(\frac{\log t - \mu}{\sigma} \right)^2 \right]}{t \int_0^\infty \frac{1}{\sigma \sqrt{2\pi}} f(t) dt}, \quad t > 0.$$

- **MTTF**

$$E(T) = \exp\left(\mu + \frac{\sigma^2}{2}\right).$$

- **La variance**

$$V(T) = \exp\left(2\mu + \frac{\sigma^2}{2}\right) (\exp(\sigma^2) - 1).$$

Remarques 3.6.1 1) Il est important de noter que si X_1 et X_2 suivent, de manière indépendante, des loi log-normales, alors la variable $Y = X_1 X_2$ suit également une loi log-normale [48].

2) Cette loi est souvent utilisée pour représenter les durées de réparation des composants ou les incertitudes dans la connaissance d'une donnée de sûreté de fonctionnement.

3) Le domaine de définition n'étant jamais négatif, il n'y a aucune limitation à l'emploi de la distribution log-normale en fiabilité. Le taux de défaillance est croissant dans le début de vie puis décroissant en tendant vers zéro et la distribution est très dissymétrique [47].

Loi de Pareto

Cette loi a été introduite pour modéliser la distribution de revenus supérieurs à un seuil donné, puis s'est avérée utile pour d'autres domaines, tels que le domaine de fiabilité (dans la méthode ABC) [31]. Elle a deux paramètres strictement positifs. Un paramètre de seuil α et un paramètre de forme t_0 . On dit qu' une variable aléatoire T suit une loi de Pareto, de paramètre α et t_0 si :

- **Densité de probabilité**

$$f(t) = \begin{cases} \frac{t_0}{\alpha} \left(\frac{\alpha}{t}\right)^{t_0+1} & \text{si } \alpha > 0 \\ 0, & \text{si } t_0 > 0. \end{cases}$$

- **Fonction de répartition**

$$F(t) = \begin{cases} 1 - \left(\frac{\alpha}{t}\right)^{t_0} & \text{si } t \geq \alpha \\ 0, & \text{si } t < \alpha. \end{cases}$$

- **Fonction de fiabilité**

$$R(t) = \left(\frac{\alpha}{t}\right)^{t_0}, \text{ si } t \geq \alpha.$$

- **Taux de défaillance**

$$\lambda(t) = \frac{t_0}{t}, \text{ si } t > t_0.$$

Remarques 3.6.2 1) La moyenne n'existe que si $t_0 > 1$ et est égale à:

$$E(T) = \left(\frac{t_0}{t_0 - 1}\right) \alpha.$$

2) La variance n'existe que si $t_0 > 2$ et est égale à:

$$V(T) = \frac{\alpha^2 t_0}{(t_0 - 2)(t_0 - 1)^2}.$$

Loi discrète

Loi de Poisson

Dans le cas où le nombre d'expériences tend vers l'infini, mais que la probabilité de d'un événement reste faible $p < 0.1$, il est préférable d'utiliser la loi de poisson au lieu de la Binomiale. La distribution de Poisson est généralement utilisée pour évaluer les risques dans l'industrie, elle prend en compte les problèmes en temps et en espace. D'une façon générale, c'est la loi qui traduit un nombre de réalisations d'événements très peu probables. Dans une suite très nombreuse d'épreuves aléatoires ($n \geq 50$), la probabilité de réalisation des événements est toujours la même.

La loi de poisson est une loi de paramètre $\lambda > 0$, défini par

$$P(T = k) = \frac{\lambda^k}{k!} \exp(-\lambda), \lambda > 0, k \in \mathbb{N}.$$

- **MTTF**

$$E(T) = \lambda.$$

- **La variance**

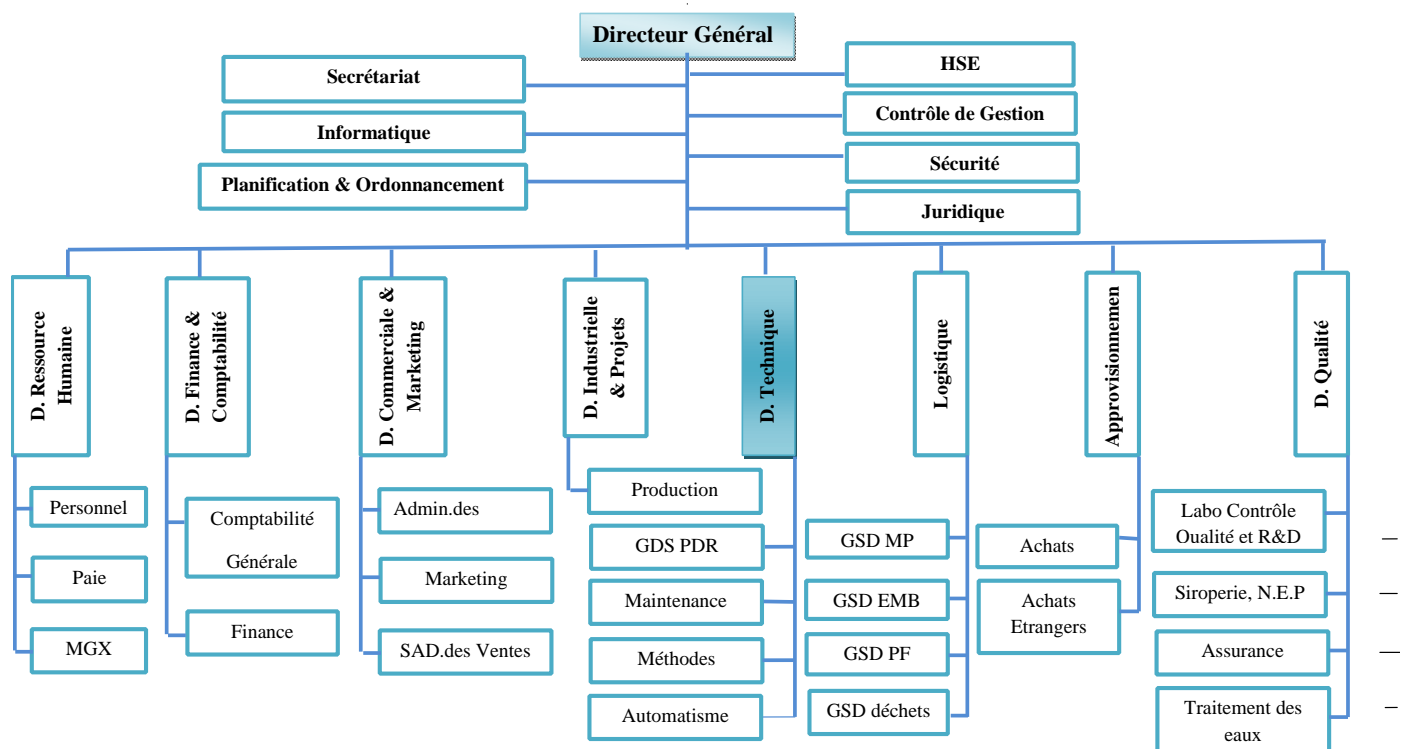
$$V(T) = \lambda.$$

Remarques 3.6.3 1) *La loi de poisson est une loi discrète dont la variable aléatoire est un nombre d'événements, alors que dans la loi Exponentielle, la variable aléatoire est l'intervalle de temps entre deux événements successifs.*

2) *En pratique, si $\lambda > 18$, la loi de poisson peut être assimilée à une loi normale de moyenne λ , et de variance λ .*

Annexe B

I) Organigramme SARL IBRAHIM & Fils Ifri



II) Description de la machine souffeuse

La souffeuse est une machine destinée à la fabrication des bouteilles vides à partir des préformes qui ont une structure de tube, fabriquées dans une autre unité spécialisée en PET. Pour atteindre ce but la souffeuse suit les étapes suivantes :

- Préchauffage des préformes ;
- Etirage axial mécanique ;
- Etirage radial par soufflage d'air comprimé à 40 bars à l'intérieur d'un moule.

La souffeuse est constituée de quatre parties principales :

- Partie d'alimentation des préformes froides ;
- Partie du four linéaire pour le préchauffage des préformes ;
- Partie pour le soufflage des préformes ;
- Partie de sortie des préformes.



Figure.1- Présentation de la machine souffleuse.

Courbe de Pareto

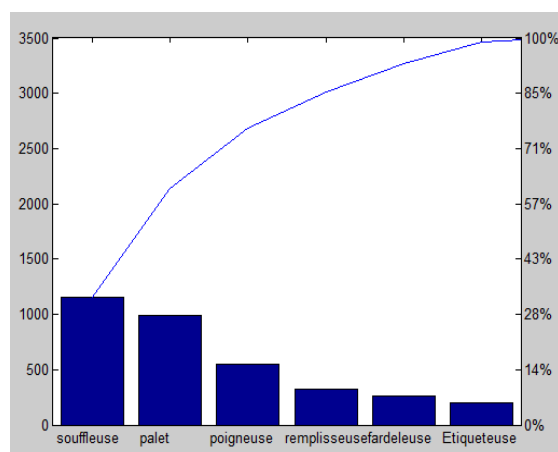


Figure.2. Courbe de Pareto pour la ligne
Combi 20.

Bibliographie

- [1] Augé J-C., « Utilisation d'un modèle de hasard proportionnel pour estimer la fiabilité de composants mécanique », Thèse de doctorat en mécanique, Ecole Centrale de Lyon, 1998.
- [2] Beaumont P ., « *Optimisation des plans d'essais accélérés application à la tenue en fatigue de pièces métalliques de liaison au sol* », Thèse de doctorat en sciences pour l'ingénieur, l'université d'Angers
- [3] Belkacem C., « *Stabilité des modèles de rupture sous contamination* », Thèse de doctorat en mathématiques, Université Mouloud-Mammeri, Tizi-Ouzou, 2013.
- [4] Belkacem N., « *Modèle d'incertitude appliqués au problème de managment de l'eau* », Mémoire de magister en mathématiques, Université M-Mammeri, Tizi-Ouzou, 2012.
- [5] Benaddy M., « *Modélisation et prédiction de la fiabilité des logiciels et des réseaux* », Thèse de doctorat en informatique, Université Ibn Zohr Agadir, 2012.
- [6] Bernier J., Boreux J-J., Parent E., « *Pratique du calcul bayésien* », Ed Springer, Paris, 2010
- [7] Bouchaib R., El Hami A., « *Fiabilité et optimisation des systèmes : Théorie et applications cours et exercices corrigés* », Ed Ellipses, 2011.
- [8] Bouhou S., Braham H., « *Application de l'inférence bayésienne en fiabilité* », Mémoire de master en Mathématiques, Université A-Mira Bejaïa, 2013.
- [9] Boumarafi M., « *Modèle de réparation minimale basé sur l'actualisation bayésienne de la probabilité de défaillance à la sollicitation des matériels de sauvegarde* », Service IPN département SID 1, Avenue du Général de Gauld, 1994.

-
- [10] Bunouf P., «*Lois bayésiennes a priori dans un plan binomial séquentiel*», Thèse de doctorat En Mathématique, Université de Rouen, 2006.
- [11] Bousquet N., «*Analyse bayésienne de la durée de vie de composants industriels*», Thèse de doctorat en science, Université Paris du Sud, 2006.
- [12] Bousquet N., «*A bayesian analysis of industrial lifetime data with Weibull distributions*», Rapport de recherche INRIA, Université d'Orsay, Novembre 2006.
- [13] Charpentier D., Lanternier B., Lyonnet P., «*Taux de défaillance bayésien pondéré à partir de données de retour d'expérience*», Ed Hal, Casablanca, Marroco, Novembre 2005.
- [14] Coccozza-Thivent C., «*Processus stochastiques et fiabilité des systèmes*, Mathématique & Application 28», Ed Springer, Verlag Berlin Heidelberg, 1997.
- [15] Craciun G., «*Fonctions de perte en actuariat* », Mémoire en vue de l'obtention du grade de maître en sciences, Université de Montréal, Septembre 2009.
- [16] Dupuis J., «*Statistique bayésienne et algorithme MCMC*», Cours IMAT, Septembre 2007.
- [17] Djafari A-M., «*Maximum Entropy and inverse problems in image reconstruction*», Synthèse en sciences, Laboratoire des signaux et systèmes, France, 1993.
- [18] Droesbek J-J., Fine J., Sporta G., «*Méthode bayésienne en statistique*», Ed Technip, Paris, 2002.
- [19] Fertou E., Procaccia M., «*Fiabilité et maintenance des matériels industriels réparables et non réparables*», Ed Lavoisier, Paris, Juillet 2011.
- [20] Ghouil D., «*Aspect de la robustesse bayésienne dans les modèles AR(1)*», Mémoire de magister en Mathématique, Université M-Mammeri, Tizi-Ouzou.
- [21] Ghouar A., «*L'estimation de la fonction de fiabilité*», Mémoire de Magistère en Mathématiques, Université Badji Mokhtar-Annaba, 2013.

-
- [22] Grous A., «*Eléments d'analyse de la fiabilité et du contrôle de qualité : statistique appliquées par exemple*», Ed Lavoisier, Paris, 2013.
- [23] Guessoum Y., «*Modélisation probabiliste de l'impact des actions de maintenance sur le comportement des matériels électronique de réseaux*», Thèse de doctorat, Spécialité : Optimisation et sûreté de fonctionnement, Université de Technologie de Troyes, 2012.
- [24] Houadi M., Moulaoui F., «*Diagnostic de la maintenance et réestimation des paramètres de fiabilité dans un environnement de données manquantes*», mémoire de master en vue du diplôme d'ingénieur d'état en recherche opérationnelle, université A-Mira Bejaïa, 2004.
- [25] Kass R.E., Wasserman L., «*The selection of prior distributions by formal rules*», Journal of Statistical Association, Vol.91, N°435, Septembre 1996.
- [26] Lagha K., «*Cours Statistique Bayésienne, master 1 Statistique et Analyse Décisionnelle*», Université de A.MIRA Bejaïa, 2014.
- [27] Laggoune R., «*Optimisation de la maintenance par la fiabilité opérationnelle des systèmes mécaniques multi-composants. Application industrielles*», Thèse de doctorat en Sciences, Université A. Mira Bejaïa.
- [28] Lambert A., «*Théorie de la mesure et intégration*» Cours de mathématiques L3, Université Pierre & Marie Curie (Paris 6), 2010.
- [29] Laronde R., «*Fiabilité et durabilité d'un système complexe dédié aux énergies renouvelables application à un système photovoltaïque*», Thèse de doctorat en sciences de l'ingénieur, Université Angers, 62 avenue notre dame du lac, 2011.
- [30] Lasnier G., «*Sûreté de fonctionnement des équipements et calcul de fiabilité*», Ed Lavoisier, Paris, 2011.
- [31] Lejeune M., «*Statistique la théorie et ses applications*», Ed Springer, N°2, Paris, 2010.
- [32] Ligeron J-C., Lionnet P., «*La fiabilité en exploitation, organisation et traitement des données*», Ed Lavoisier, Paris, 1993.

-
- [33] Lionnet P., «*Optimisation d'une politique de maintenance*», Ed Lavoisier, Paris, 1993.
- [34] Marin J-M., Robert C P., «*Bayesian core A practical to computational Bayesian statistics*», Ed Springer, Paris, 2007.
- [35] Mihalache A G., «*Modélisation et évaluation de la fiabilité des systèmes mécaniques : application sur système Embarque*», Thèse de doctorat en sciences de l'ingénieur, Université d'Angers, 2007.
- [36] Mouy M., «*Estimation de la fiabilité d'un système et de ses composants*», Mémoire de magister en mathématique, Université de Mentouri, 2009.
- [37] Parent E., Bernier J., «*Le raisonnement bayésien modélisation et inférence modélisation et inférence*», Ed Springer, Paris, 2007.
- [38] Philip A., «*Statistique Bayésienne*», Laboratoire de mathématiques Jean Leray, Université Paris XI, 2006.
- [39] Procaccia H., «*Introduction à l'analyse probabiliste des risques industriels*», Lavoisier, Paris, février 2009.
- [40] Procaccia H., «*Evaluation bayésienne de la probabilité de défaillance à la sollicitation des matériels de sauvegarde*», Rapport de recherche, Avenue du Général-de-Gaulle, 1994.
- [41] Piepszownik L., Procaccia H., «*Fiabilité des équipements et Théorie de la décision statistique fréquentielle et bayésienne*», Ed Eyrolles, Collection de la Direction des Etudes et Recherches d'Electricité de France, 1992.
- [42] Ringler J., «*Une modélisation bayésienne du taux de défaillance en fiabilité*», revue de Statistique Appliquée, tome 29, N°1, P43-56, 1981.
- [43] Robert C P., «*Le choix bayésien Principes et pratique*», Ed Springer-Verlag France, Paris, 2006.
- [44] Rousseau J., «*Statistique bayésienne*», notes de cours, 3eme années Statistique, Université D'auphine, Paris, 2009-2010.

- [45] Sporta G., «*Probabilité, Analyse Des Données et Statistiques*», Ed Technique, Paris, 1990.
- [46] Suhner M C., Procaccia H., «*Démarche bayésienne et application à la sûreté de fonctionnement* », Ed Lavoisier, France, janvier 2003.
- [47] Tebbi O., «*Estimations des lois de fiabilité en mécanique par les essais accélérés* », Thèse de doctorat en sciences de l'ingénieur, Université d'Angers, Mars 2005.
- [48] Villemeur A., «*Sûreté de fonctionnement des systèmes industrielles, Fiabilité-Facteur humain-informatisation*», Ed Eyrolles, Paris, 1988.
- [49] Zhu Q ., «*Modeles bayésiens et application a l'estimation des caractéristiques de produits finis et au contrôle de la qualité*», Thèse de doctorat , Paris, juillet 1991.

Résumé

L'analyse de fiabilité des équipements dans le domaine industriel se fait sur la base des méthodes statistiques. Dans ce travail, nous avons utilisé l'approche bayésienne dans le domaine de fiabilité qui est l'une des techniques les plus explorées à l'heure actuelle, permettant de combiner les données objectives (retour d'expérience) et les données subjectives (avis d'experts et essais antérieurs sur des matériels identiques). De ce fait, nous avons présenté les différentes méthodes de construction de la loi a priori dans le cas informatif et non informatif. Deux applications en fiabilité, ont été présentées pour l'estimation des paramètres de fiabilité donnés par: la probabilité de défaillance à la sollicitation, le taux de défaillance et le temps moyen de bon fonctionnement. Une autre application avec les données du stage, effectué à l'entreprise IFRI BEJAIA, a révélé l'impact de l'information a priori sur l'estimation des paramètres de fiabilité.

Mot clés: Statistique bayésienne, Fiabilité, Information a priori, Taux de défaillance.

Abstract

The reliability analysis in the industrial domain is based on statistical methods. In this work, we used the bayesian approach in the reliability which combines the objective data (the experience return) with the subjective data (opinions of experts and previous tests on the identical materials). For this, we presente different methods of construction of the prior informative and non informative case. We give applications in reliability to estimate the parameters of reliability with bayesian approach. Another application is given with training data of IFRI BEJAIA company where the impact of prior information is proved on the estimation.

Word keys: bayesian approach, Reliability, prior Information, failure rate.