

RÉPUBLIQUE ALGÉRIENNE DÉMOCRATIQUE ET POPULAIRE
MINISTÈRE DE L'ENSEIGNEMENT SUPÉRIEUR ET DE LA RECHERCHE SCIENTIFIQUE
UNIVERSITÉ MOULOUD MAAMERI DE TIZI-OUZOU
FACULTÉ DES SCIENCES
DÉPARTEMENT DE MATHÉMATIQUES



MÉMOIRE DE FIN DE CYCLE EN VUE D'OBTENTION DU DIPLÔME DE MASTER

OPTION : PROBABILITÉS ET STATISTIQUE

Inférence des paramètres de la Loi Gamma sous diverses approches, Application à la fiabilité.

Thème présenté par **MLLE MENCHAR SARAH**

Devant le jury composé de :

M. FELLAG HOCINE :	Professeur	UMMTO	- Président
Mme. ATIL LYNDA :	MCA	UMMTO	- Examinatrice
Mme. BELKACEM CHERIFA :	MCB	UMMTO	- Rapporteur

Promotion 2022-2023

Je tiens à exprimer ma gratitude envers mon encadrante Madame Belkacem, pour le soutien et l'encadrement qu'elle a pu m'offrir durant la préparation de ce mémoire. Son expertise, sa patience et son dévouement m'ont été d'une grande aide.

Je tiens aussi à remercier Monsieur Fellag et Madame Atil pour avoir accepté de faire partie du jury et pour leur précieuse contribution à l'évaluation de mon mémoire.

Je remerci aussi mes enseignants et toute l'équipe pédagogique du Master Probabilité-Statistiques, grâce à vos conseils avisés et à votre guidance, j'ai pu acquérir de nouvelles compétences et approfondir mes connaissances. Votre accompagnement m'a permis de progresser rapidement et efficacement, et je suis reconnaissante pour tout ce que vous avez pu faire pour moi et mes camarades.

Encore une fois, je vous remercie sincèrement et je tiens à vous exprimer ma plus profonde reconnaissance. Je suis fière et honorée d'avoir eu la chance de travailler avec vous.

« La plus grande récompense qu'un homme obtienne pour son labeur n'est pas ce qu'il en a retiré, mais en quoi cela l'a transformé. »

-John Ruskin

Je dédie ce modeste travail tout d'abord à mes parents, la prunelle de mes yeux. Ils ont été non seulement des parents aimants, mais aussi des amis fidèles et des mentors inspirants. Leur dévouement envers moi a été incroyablement fort et j'ai toujours pu compter sur leur soutien et leur amour inconditionnel. Ils m'ont aidée à grandir, à apprendre et à me développer en la personne que je suis aujourd'hui.

Ils ont travaillé dur pour me donner les meilleures opportunités, pour me soutenir dans mes choix et pour me guider sur le chemin de la réussite académique. Je suis fière de pouvoir célébrer cette étape importante de ma vie en leur honneur et en reconnaissance de tous les sacrifices qu'ils ont pu faire pour moi.

Puis à tous mes proches, mes amies, mes camarades de promo, mes frères Ahmed et Massi, mon meilleur ami Kamal et ma meilleure amie Doudouche. Leur soutien tout au long de cette année était très précieux pour moi.

Et enfin, une pieuse pensée pour mes grand-mères, Azouzou et Jidda, les femmes exceptionnelles qui ont illuminé nos vies de leur gentillesse, de leur sagesse et de leur amour. Leur mémoire perdurera en moi pour toujours.

TABLE DES MATIÈRES

1	Généralités sur la Loi Gamma	12
1.1	Introduction	12
1.1.1	Fonction de répartition	13
1.1.2	Fonction de densité	14
1.1.3	Fonction de survie	15
1.1.4	Fonction de hasard	15
1.1.5	Moments d'ordre k	16
1.1.6	Fonction génératrice des moments	17
1.1.7	Fonction caractéristique	18
1.1.8	Simulation d'une loi Gamma	18
1.2	Liens avec d'autres lois	19
1.2.1	Loi Exponentielle	19
1.2.2	Loi du Chi Carré χ^2	19
1.2.3	Distribution d'Erlang	20
1.2.4	Distribution de Dirichlet en Forme de Point	20
1.2.5	Distribution Bêta	20
1.3	Somme de deux lois Gamma	20
1.4	L'Inverse Gamma	21
1.4.1	Les moments d'ordre k	21

1.5	La loi Gamma généralisée	22
1.6	Application de la loi gamma	22
2	Estimation des paramètres de la loi Gamma	25
2.1	Estimation Classique	25
2.1.1	Méthode des Moments	25
2.1.2	Méthode du Maximum de vraisemblance	31
2.2	Estimation Bayésienne	35
2.2.1	Modèle Bayésien	35
2.2.2	Théorème de Bayes	37
2.2.3	Fonctions coût usuelles	38
2.2.4	Estimateur de Bayes	40
2.3	Estimation Bayésienne des paramètres de la loi Gamma	41
2.3.1	Les estimateurs Bayésiens sous différentes fonctions coût	42
3	Application de la loi Gamma à la fiabilité	50
3.1	Introduction	50
3.2	Notions générales	51
3.3	Mesures de fiabilité	52
3.3.1	Mesures pour systèmes non réparables	52
3.3.2	Mesures pour systèmes réparables	53
3.3.3	Lois classiques utilisées dans la fiabilité	55
3.4	Censure	55
3.4.1	Type I	56
3.4.2	Type II	57
3.5	Application de la loi Gamma à la fiabilité	57
3.5.1	Application sur des données réelles	58
3.6	Conclusion	65
.1	La Méthode de Lindley	67

TABLE DES FIGURES

1.1	Graphe d'une répartition d'une loi Gamma	13
1.2	Graphes de deux densités de loi Gamma	14
1.3	Graphe d'une fonction de survie d'une loi Gamma	15
1.4	Graphe d'une fonction de hasard d'une loi Gamma	16
1.5	Histogramme des échantillons d'une loi Gamma(1,3)	19
1.6	Graphe d'une densité d'une loi Gamma inverse	21
3.1	Taux de défaillance d'une loi Gamma	56

LISTE DES TABLEAUX

2.1	Quelques estimateurs de Bayes usuels.	40
3.1	Données de fiabilité de voyants lumineux d'avion	59
3.2	Données pour les exemples simulés	64
3.3	Données pour les exemples simulés	64
3.4	Résultats des exemples	65

RÉSUMÉ

Ce mémoire examine en profondeur la loi gamma, une distribution de probabilité continue couramment utilisée en statistiques et en probabilités. Elle explore les propriétés mathématiques de cette distribution, notamment les paramètres de forme et d'échelle, ainsi que la manière dont elle peut être appliquée pour modéliser divers phénomènes dans des domaines tels que la fiabilité des systèmes. Il offre ainsi un aperçu complet de l'importance et de la polyvalence de la loi gamma dans la modélisation statistique et propose des contributions significatives à la compréhension et à l'application de cette distribution dans différents domaines de recherche et d'industrie.

ABSTRACT

This thesis thoroughly examines the gamma distribution, a continuous probability distribution commonly used in statistics and probability theory. It explores the mathematical properties of this distribution, including shape and scale parameters, as well as how it can be applied to model various phenomena in fields such as system reliability. It provides a comprehensive overview of the significance and versatility of the gamma distribution in statistical modelling and makes significant contributions to the understanding and application of this distribution in various research and industry domains.

INTRODUCTION GÉNÉRALE

La loi gamma est une distribution de probabilité continue qui trouve une grande importance dans un large éventail de domaines, de la statistique à l'ingénierie, en passant par les sciences sociales et la médecine. Elle est souvent utilisée pour modéliser des événements aléatoires qui se produisent dans le temps ou l'espace.

L'histoire de la loi gamma remonte au 18^e siècle avec le développement de la théorie des probabilités. Daniel Bernoulli, un mathématicien suisse en 1738, est crédité d'avoir introduit la fonction gamma dans son ouvrage "Ars Conjectandi". Cependant, il n'a pas utilisé explicitement le terme "loi gamma". En (1812) : le mathématicien et statisticien français Laplace a joué un rôle clé dans le développement des distributions de probabilité. Dans son ouvrage "Théorie analytique des probabilités", il a introduit la notion de "loi de probabilité" et a utilisé la fonction gamma pour décrire la distribution des erreurs, ensuite en 1895 le statisticien britannique Pearson qui a été l'un des premiers à définir formellement la loi gamma et à la nommer, a introduit la fonction de densité de probabilité de cette loi et a étudié ses propriétés statistiques, enfin en 1924 le statisticien britannique Fisher, a joué un rôle essentiel dans l'établissement de la méthode de maximum de vraisemblance et son application à la loi gamma. Il a développé des techniques pour estimer les paramètres de cette loi à partir de données observées.

La distribution gamma présente des propriétés mathématiques intéressantes, notamment une flexibilité dans la modélisation de la variabilité des données, ce qui en fait un outil puissant pour l'analyse statistique.

L'objectif de ce mémoire est d'explorer en profondeur la distribution gamma. Dans le premier chapitre nous examinerons ses caractéristiques, ses propriétés,

ses extensions, nous parlerons sur les aspects théoriques de cette distribution, en étudiant ses moments, sa fonction de densité de probabilité.

Dans le deuxième chapitre, nous nous pencherons d'avantage sur les méthodes d'estimation des paramètres de la distribution.

Enfin, dans le dernier chapitre, nous aborderons un exemple d'application pratiques de la loi gamma dans la fiabilité, tels que la modélisation des temps de défaillance dans l'industrie où la modélisation des données est cruciale et nous terminerons par une conclusion générale et quelques perspectives.

CHAPITRE 1

GÉNÉRALITÉS SUR LA LOI GAMMA

1.1 Introduction

La loi gamma est l'une des distributions de probabilité les plus importantes et les plus polyvalentes en statistiques et en probabilités. Elle trouve des applications dans de nombreux domaines, notamment en sciences naturelles, en ingénierie, en économie, en biostatistique et bien d'autres encore. Cette distribution tire son nom du fait qu'elle est étroitement liée à la fonction gamma, une fonction mathématique cruciale dans son développement.

La distribution Gamma est particulièrement adaptée pour modéliser des variables aléatoires positives, continues et à valeurs réelles. Elle est souvent utilisée pour représenter des temps de défaillance, des durées, des intervalles entre événements, ou toute autre quantité positive qui suit une distribution continue. Elle est caractérisée par deux paramètres : le paramètre de forme (alpha) et le paramètre d'échelle (beta).

La loi gamma est souvent utilisée en conjonction avec d'autres lois de probabilité pour modéliser des phénomènes plus complexes. Elle est également utilisée dans les modèles de régression pour ajuster des données expérimentales et pour estimer des paramètres inconnus dans les modèles statistiques.

1.1.1 Fonction de répartition

La fonction de répartition "F" de la loi gamma s'écrit comme suit :

$$\forall x > 0, F(x; \alpha, \beta) = \frac{\gamma(\alpha, x)}{\Gamma(\alpha)} \quad (1.1)$$

Où

1. $\gamma(\alpha, x)$ est la fonction gamma incomplète régulière, qui est utilisée pour calculer la probabilité que la variable aléatoire gamma soit supérieure à x , elle est définie comme suit :

$$\gamma(\alpha, x) = \int_0^x t^{\alpha-1} e^{-t} dt$$

2. Γ est la fonction eulérienne définie par l'intégrale pour $x > 0$ par

$$\Gamma(\alpha) = \int_0^{\infty} t^{\alpha-1} e^{-t} dt$$

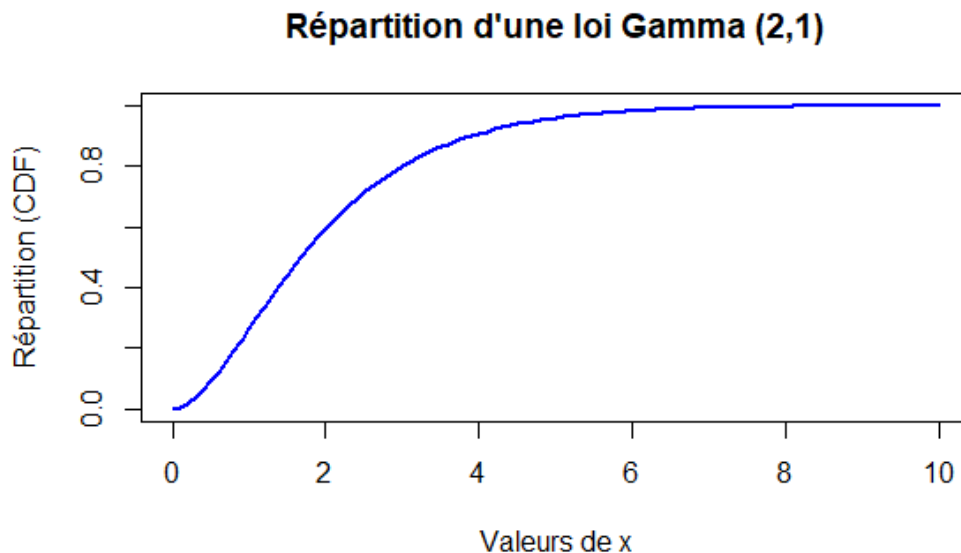


FIGURE 1.1 – Graphe d'une répartition d'une loi Gamma

1.1.2 Fonction de densité

La fonction de densité "f" de la loi gamma s'écrit comme suit :

$$\forall x > 0, f(x; \alpha, \beta) = \frac{\beta^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-\beta x} \quad (1.2)$$

Où

1. x est la variable aléatoire continue qui suit une loi gamma avec les paramètres α et β
2. α est le paramètre de forme (strictement positif)
3. β est le paramètre d'échelle (strictement positif)

Remarque Le paramètre de forme détermine la forme de la distribution et le paramètre d'échelle détermine l'étendue de la distribution. Plus spécifiquement, la loi gamma a une forme en cloche, avec une asymétrie à droite si le paramètre de forme est inférieur à 1 et une asymétrie à gauche si le paramètre de forme est supérieur à 1.

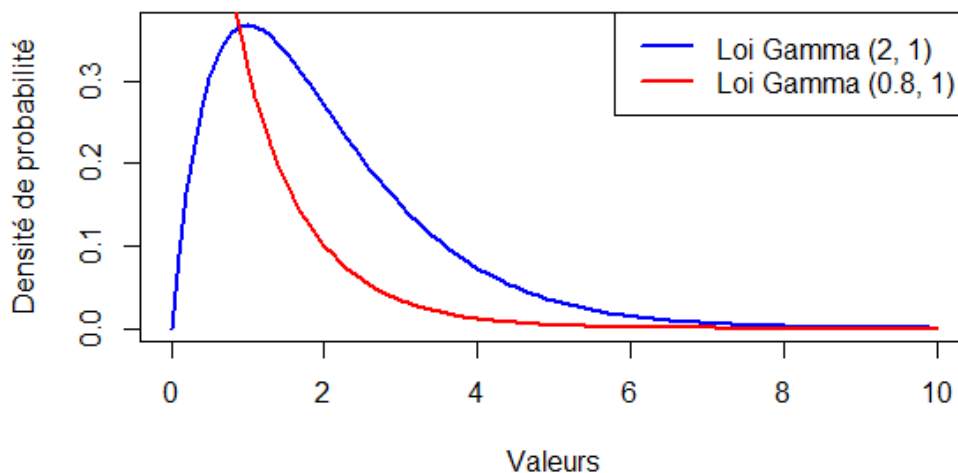


FIGURE 1.2 – Graphes de deux densités de loi Gamma

1.1.3 Fonction de survie

La fonction de survie de la loi gamma est définie comme suit :

$$S(x; \alpha, \beta) = 1 - \frac{\gamma(\alpha, x)}{\Gamma(\alpha)} \quad (1.3)$$

Où x est la valeur à laquelle vous souhaitez calculer la fonction de survie.

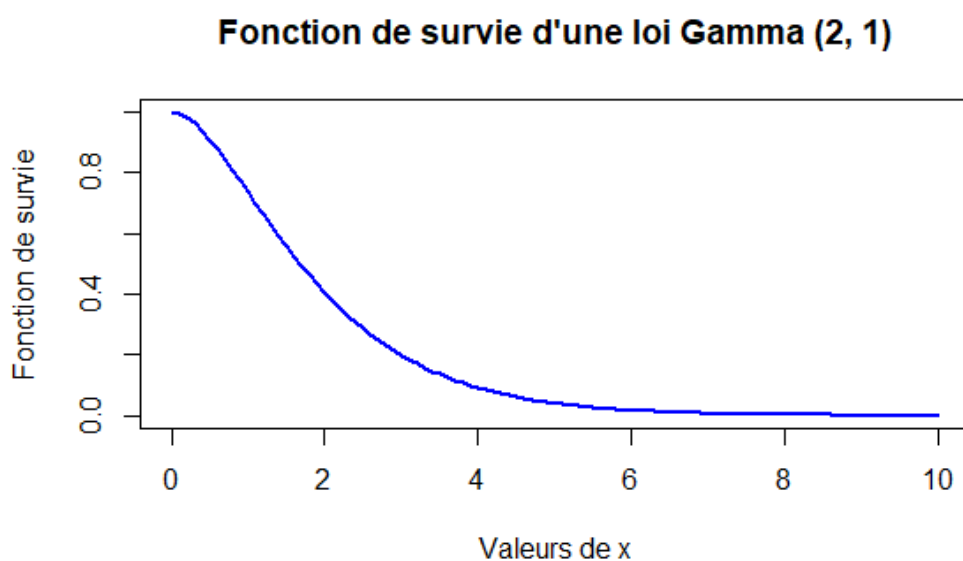


FIGURE 1.3 – Graphe d'une fonction de survie d'une loi Gamma

1.1.4 Fonction de hasard

La fonction de hasard (hazard function) d'une loi Gamma est définie comme suit :

$$h(x; \alpha, \beta) = \frac{S(x; \alpha, \beta)}{f(x; \alpha, \beta)} \quad (1.4)$$

Où

1. x est la valeur à laquelle nous souhaitons calculer la fonction de hasard.
2. $f(x; \alpha, \beta)$ est la densité de probabilité.
3. $S(x; \alpha, \beta)$ est la fonction de survie.

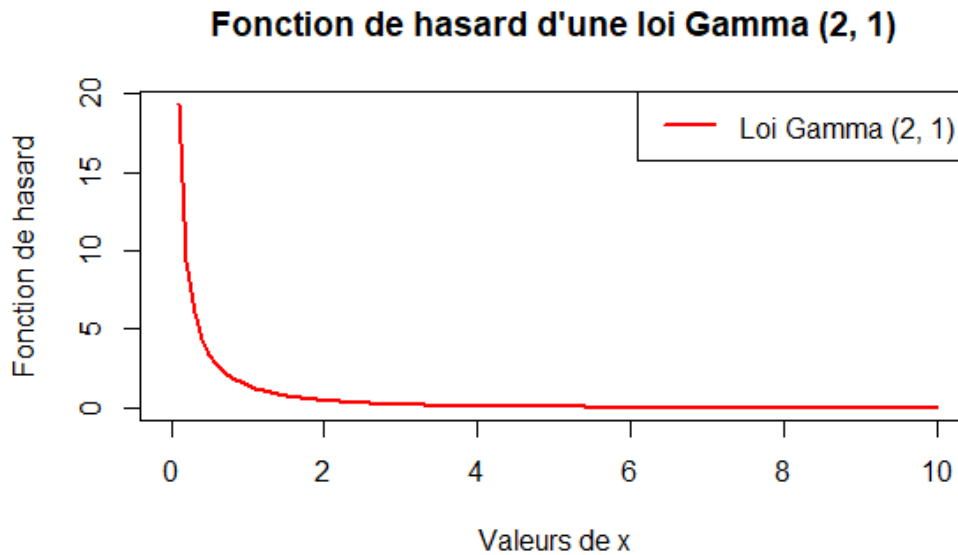


FIGURE 1.4 – Graphe d’une fonction de hasard d’une loi Gamma

1.1.5 Moments d’ordre k

Si X est une variable aléatoire admettant une densité f , alors le moment d’ordre k se calcule de la façon suivante :

$$\mu_k = \int_{\mathbb{R}} x^k f(x) d(x) \quad (1.5)$$

Espérance

L’espérance est le moment d’ordre 1, elle est calculée comme suit :

$$\begin{aligned} \mathbb{E}(X) &= \int_0^{\infty} x f(x; \alpha, \beta) dx \\ &= \int_0^{\infty} \frac{\beta^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-\beta x} dx \\ &= \frac{\beta^\alpha}{\Gamma(\alpha)} \int_0^{\infty} x^{\alpha-1} e^{-\beta x} dx \end{aligned}$$

Nous posons le changement de variable $x = \beta y$, nous obtenons :

$$\begin{aligned}\mathbb{E}(X) &= \frac{\beta^\alpha}{\Gamma(\alpha)} \int_0^\infty e^{-y} \left(\frac{y}{\beta}\right) dy \\ &= \frac{\beta^\alpha}{\Gamma(\alpha)} \frac{\Gamma(\alpha + 1)}{\beta^{\alpha+1}}\end{aligned}$$

Ce qui nous donne :

$$\mathbb{E}(X) = \frac{\alpha}{\beta} \quad (1.6)$$

Variance

Nous savons que :

$$\text{Var}(X) = \mathbb{E}(X^2) - (\mathbb{E}(X))^2 \quad (1.7)$$

Calculons d'abord le moment d'ordre $\mathbb{E}(X)^2$,

Par le même changement de variable :

$$\begin{aligned}\mathbb{E}(X^2) &= \int_0^\infty x^2 f(x; \alpha, \beta) dx \\ &= \frac{\beta^\alpha}{\beta^{\alpha+2}\Gamma(\alpha)} \int_0^\infty e^{-y} y^{\alpha+1} dy \\ &= \frac{\Gamma(\alpha + 2)}{\beta^2\Gamma(\alpha)}\end{aligned}$$

Ce qui nous donne :

$$\mathbb{E}(X^2) = \frac{\alpha(\alpha + 1)}{\beta^2} \quad (1.8)$$

En remplaçant (1.6) et (1.8) dans (1.7), nous aurons

$$\text{Var}(X) = \frac{\alpha}{\beta^2} \quad (1.9)$$

1.1.6 Fonction génératrice des moments

La fonction génératrice est utilisée en théorie des probabilités pour déterminer les moments de la distribution. Elle est définie comme suit :

$$g_x(t) = \mathbb{E}(e^{tx}) = \left(1 - \frac{t}{\beta}\right)^{-\alpha} \quad (1.10)$$

Où :

1. t est le paramètre de la fonction génératrice.
2. X est la variable aléatoire qui suit la loi gamma.

1.1.7 Fonction caractéristique

La fonction caractéristique est une transformée de Fourier complexe de la densité de probabilité de la variable aléatoire. Elle permet de caractériser complètement la distribution de la variable aléatoire. Elle est définie par :

$$\varphi_x(t) = \left(\frac{\beta}{\beta - it} \right)^\alpha \quad (1.11)$$

Où :

1. i est la "unit imaginary unit" (unité imaginaire).
2. t est le paramètre de la fonction caractéristique.

1.1.8 Simulation d'une loi Gamma

Pour simuler une variable aléatoire suivant une loi gamma en utilisant R, nous pouvons utiliser la fonction `rgamma()`. Cette fonction génère des échantillons aléatoires à partir d'une distribution gamma avec des paramètres spécifiés.

Voici un exemple de programme pour simuler une loi gamma :

```
# Paramètres de la loi Gamma
alpha = 1 # Paramètre de forme
beta = 3  # Paramètre d'échelle

# Nombre d'échantillons à générer
n = 1000

# Génération d'échantillons aléatoires suivant une loi Gamma
echantillons = rgamma(n, shape = alpha, scale = beta)

# Affichage des premiers échantillons
head(echantillons)
[1] 1.1759368 4.7204821 1.3860614
     0.1483395 4.6556286 1.6782298

# Tracer un histogramme des échantillons
hist(echantillons, breaks = 30, main = "Histogramme de la loi
      xlab = "Valeurs", ylab = "Fréquence", col = "lightblue")
```

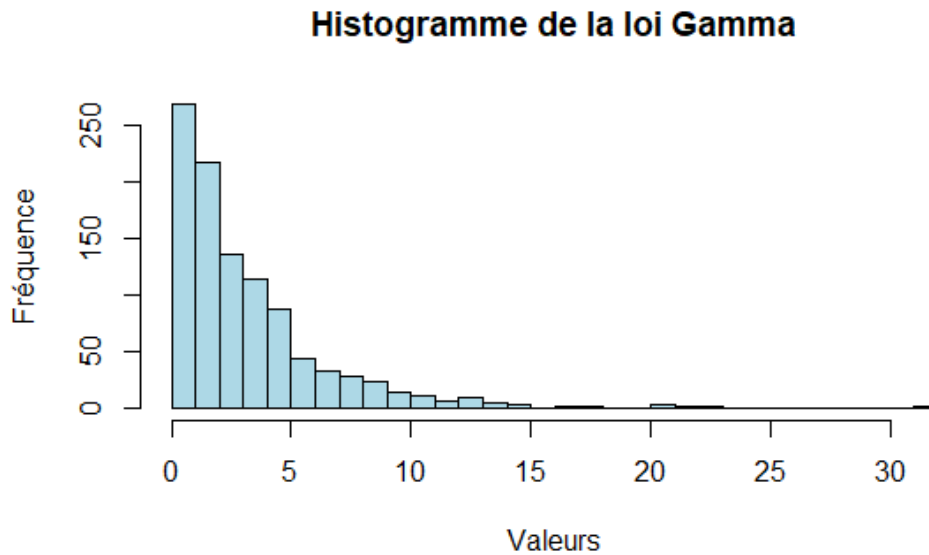


FIGURE 1.5 – Histogramme des échantillons d'une loi Gamma(1,3)

1.2 Liens avec d'autres lois

La loi Gamma est liée à plusieurs autres distributions statistiques importantes, les voici :

1.2.1 Loi Exponentielle

Si X suit une loi gamma de paramètres $\alpha = 1$ et $\beta = 1$, alors X suit une loi exponentielle de paramètre λ .

1.2.2 Loi du Chi Carré χ^2

La distribution gamma est une généralisation de la distribution du chi carré. Si X suit une loi gamma avec des paramètres $\alpha = k/2$ et $\beta = 2$, alors $2X$ suit une loi du χ^2 à k degrés de liberté (k est un nombre entier positif).

1.2.3 Distribution d'Erlang

La distribution d'Erlang est une version spécifique de la distribution gamma, où le paramètre de forme α est un entier positif. Si X suit une loi gamma avec un paramètre de forme α étant un entier positif et un paramètre d'échelle β , alors X suit une distribution d'Erlang.

1.2.4 Distribution de Dirichlet en Forme de Point

: La distribution de Dirichlet est une distribution continue multivariée, souvent utilisée en statistiques bayésiennes pour modéliser des proportions. La distribution de Dirichlet en forme de point est une version particulière de la distribution de Dirichlet où tous les paramètres sont des nombres entiers. Elle est liée à la distribution gamma car les paramètres de la distribution de Dirichlet en forme de point sont souvent exprimés en termes de variables gamma indépendantes.

1.2.5 Distribution Bêta

La distribution bêta est liée à la distribution gamma via la distribution de Dirichlet. Si X suit une loi gamma avec un paramètre de forme α et Y suit une loi gamma avec un paramètre de forme β , alors la proportion $X/(X + Y)$ suit une distribution bêta avec des paramètres α et β .

1.3 Somme de deux lois Gamma

Soient X et Y deux variables aléatoires de loi respective $\Gamma(\alpha_1, \beta)$ et $\Gamma(\alpha_2, \beta)$ avec X et Y indépendantes.

Soit $Z = X + Y$, la variable aléatoire somme.

X et Y prennent toutes les valeurs réelles positives, Z prend donc aussi toutes les valeurs réelles positives.

La somme de deux lois Gamma indépendantes $\Gamma(\alpha_1, \beta)$ et $\Gamma(\alpha_2, \beta)$ suit la loi Gamma $\Gamma(\alpha_1 + \alpha_2, \beta)$. La démonstration a été faite dans [11].

Remarques

- Quand les variables aléatoires suivent la même loi gamma (alpha, beta), alors la somme de ces variables aléatoires suit une loi gamma (n alpha, beta).

$X_1 \sim \Gamma(\alpha, \beta), X_2 \sim \Gamma(\alpha, \beta), \dots, X_n \sim \Gamma(\alpha, \beta)$ alors $\sum_{i=1}^n X_i \sim \Gamma(n\alpha, \beta)$

-
- La somme est une Gamma uniquement quand le deuxième paramètre est constant.

1.4 L'Inverse Gamma

Si $X \sim \Gamma(\alpha, \beta)$, alors nous posons $Y = X^{-1} \sim IG(\alpha, \beta)$ (IG=Inverse-Gamma).
Ce qui nous donne la fonction de densité de l'inverse gamma :

$$f(y, \alpha, \beta) = \frac{\beta^\alpha}{\Gamma(\alpha)} y^{-\alpha-1} e^{-\frac{\beta}{y}} \quad y \geq 0 \quad (1.12)$$

Où :

- $y = \frac{1}{x}$
- α est un paramètre de forme.
- β est un paramètre d'intensité.

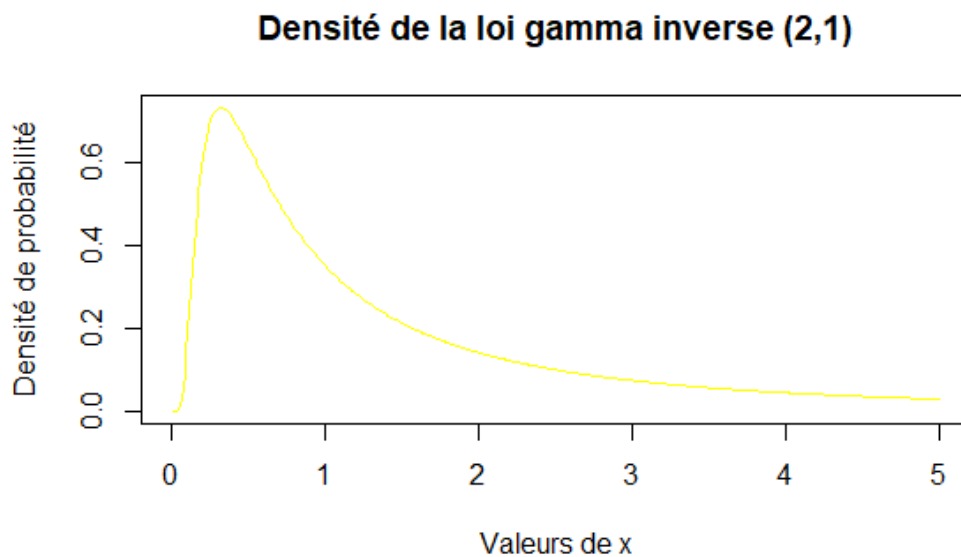


FIGURE 1.6 – Graphe d'une densité d'une loi Gamma inverse

1.4.1 Les moments d'ordre k

De la même manière que nous avons calculé l'espérance et la variance de la loi Gamma, nous aurons :

L'espérance

$$\mathbb{E}(X^{-1}) = \mathbb{E}(Y) \frac{\beta}{\alpha - 1}, \quad \alpha \geq 1 \quad (1.13)$$

La Variance

$$\text{Var}(X^{-1}) = \text{Var}(Y) = \frac{\beta^2}{(\alpha - 2)(\alpha - 1)^2} \quad (1.14)$$

1.5 La loi Gamma généralisée

La loi gamma généralisée est une distribution de probabilité continue qui est une généralisation de la distribution gamma standard. Elle est définie comme suit :

$$f(x; \alpha, \beta, d) = \frac{d}{\beta^\alpha \Gamma(\alpha/d)} \cdot x^{\alpha-1} e^{(-x/\beta)^d} \quad (1.15)$$

où :

- x est la variable aléatoire,
- α est le paramètre de forme (strictement positif),
- β est le paramètre d'échelle (strictement positif),
- d est un paramètre de forme supplémentaire (strictement positif),
- Γ est la fonction gamma.

La notation $X \sim GGD(\alpha, \beta, d)$ signifie que la variable aléatoire X suit une distribution gamma généralisée avec les mêmes paramètres.

1.6 Application de la loi gamma

La loi gamma est largement utilisée dans de nombreux domaines pour modéliser des phénomènes continus et pour prédire des résultats futurs. Voici quelques exemples de domaines d'application de la loi gamma :

- Statistique : la loi gamma est utilisée dans diverses méthodes statistiques, telles que l'analyse de variance (ANOVA), la régression linéaire, l'analyse de survie, etc.

-
- Économie et finance : la loi gamma est souvent utilisée pour modéliser les fluctuations des prix des actifs financiers, tels que les actions, les devises, les matières premières, etc. Elle est également utilisée pour modéliser les durées de vie des options et des contrats à terme. En utilisant la loi gamma, les investisseurs peuvent estimer le risque associé à ces actifs financiers.
 - Fiabilité et durée de vie : La loi gamma est souvent utilisée pour modéliser la durée de vie des produits. Par exemple, pour modéliser la durée de vie des composants électroniques, des machines, des véhicules, etc. En utilisant des modèles basés sur la loi gamma, les ingénieurs peuvent évaluer la fiabilité d'un produit et prédire le moment où il pourrait échouer.
 - Temps d'attente dans les files d'attente : La loi gamma est également utilisée pour modéliser les temps d'attente dans les files d'attente. Par exemple, elle peut être utilisée pour modéliser le temps d'attente pour obtenir un billet de cinéma, le temps d'attente pour accéder à un guichet de service clientèle, etc.
 - Temps de réponse : La loi gamma peut également être utilisée pour modéliser le temps de réponse d'un système. Par exemple, pour modéliser le temps de réponse d'un système de communication, le temps de réponse d'un système de traitement de données, etc.
 - Modèles de simulation : La loi gamma est souvent utilisée pour générer des nombres aléatoires dans les modèles de simulation. Elle peut être utilisée pour simuler le temps qu'il faut pour qu'un événement se produise, le temps qu'il faut pour qu'un utilisateur effectue une action, etc.
 - Ingénierie : pour modéliser la durée de vie de composants électroniques, la fiabilité des systèmes, la capacité des canalisations, etc.
 - Biologie : la loi gamma est souvent utilisée pour modéliser les temps de survie des organismes, tels que les animaux et les plantes. Elle est également utilisée pour modéliser les temps de réponse des systèmes biologiques, tels que les temps de digestion des aliments, les temps de guérison des blessures, etc. En utilisant la loi gamma, les biologistes peuvent comprendre les processus biologiques et les phénomènes de la nature
 - Physique : pour modéliser les temps de déclin des particules radioactives, la distribution de tailles des étoiles, la distribution de la durée des flashes lumineux, etc.

-
- Télécommunications : la loi gamma est souvent utilisée pour modéliser les temps de réponse des réseaux informatiques, les temps d'attente des appels téléphoniques, et les temps de traitement des messages. En utilisant la loi gamma, les ingénieurs peuvent optimiser les performances des réseaux et des systèmes de communication.
 - En outre, la loi gamma est également utilisée pour modéliser des processus stochastiques tels que les files d'attente, les processus de Markov et les processus de Poisson. La grande flexibilité de la loi gamma lui permet d'être appliquée dans de nombreux domaines pour modéliser une variété de phénomènes continus.

CHAPITRE 2

ESTIMATION DES PARAMÈTRES DE LA LOI GAMMA

L'estimation de la loi gamma revêt un intérêt fondamental dans de nombreux domaines, grâce à sa capacité à modéliser des phénomènes aléatoires de durée, de délais, et de fiabilité. Il y'a plusieurs méthodes pour estimer les paramètres de la distribution gamma, nous allons aborder quelques-unes dans ce chapitre.

2.1 Estimation Classique

2.1.1 Méthode des Moments

L'estimateur des moments de la loi Gamma, a été discuté par plusieurs auteurs tels que Damselth 1975 et Miller 1980.

Avant de passer à l'estimation, rappelons d'abord ce qu'est la méthode des moments :

Méthode des moments La méthode des moments consiste à trouver une fonction m , inversible et continue, et une fonction h telle que $E[|h(X)|] < \infty$ et $m(\theta) = E_{\theta}[h(X_i)]$ pour $\theta \in \Theta$.

Définition 1. L'estimateur des moments de θ , noté $\hat{\theta}_M$ est défini par :

$$\hat{\theta}_M = m^{-1} \left(\frac{1}{n} \sum_{i=1}^n h(X_i) \right)$$

Souvent, on prend $h(X_i) = x_i^r$, m correspond au moment d'ordre r de la variable X_i dont x_i est une réalisation. Autrement dit, cette méthode consiste à égaliser les moments théoriques aux moments empiriques de l'échantillon.

Le cas de la loi gamma Supposons que X est une variable aléatoire suivant une loi Gamma.

Soient (x_1, x_2, \dots, x_n) un échantillon de taille n qui sont les réalisations de (X_1, X_2, \dots, X_n) , les moments d'ordre k sont donnés par :

$$m'_r = \frac{1}{n} \sum_{i=1}^n x_i^r, r = 1, 2, \dots, k \quad (2.1)$$

Les espérances d'ordre r sont données par

$$\mu'_r = E(X^r), r = 1, 2, \dots, k \quad (2.2)$$

Nous avons :

$$m_1 = \frac{\sum_{i=1}^n x_i}{n} = \bar{x}$$

$$m_2 = \frac{\sum_{i=1}^n x_i^2}{n}$$

$$\mu'_1 = E(X) = \frac{\alpha}{\beta}$$

$$\mu'_2 = E(X^2) = \frac{\alpha}{\beta^2} + \left(\frac{\alpha}{\beta}\right)^2$$

En appliquant la méthode des moments, les estimateurs $\hat{\alpha}$ et $\hat{\beta}$ sont les estimateurs de α et β respectivement sont les solutions du système d'équations :

$$\begin{cases} m_1 = \mu_1 \\ m_2 = \mu_2 \end{cases}$$

Ce qui nous donne :

$$\hat{\alpha} = \frac{n\bar{X}^2}{\sum_{i=1}^n X_i^2 - n\bar{X}^2} \quad (2.3)$$

$$\hat{\beta} = \frac{n\bar{X}}{\sum_{i=1}^n X_i^2 - n\bar{X}^2} \quad (2.4)$$

Propriété 1. *Il est constaté que les estimateurs sont convergents mais ne sont pas efficaces.*

Pour remédier à ce problème d'efficacité, Zhu Xiaojun et al. [30] ont proposé deux nouveaux estimateurs des moments modifiés et des méthodes de réduction de biais dans leur article.

Estimateur des moments modifiés MMe

Pour trouver le nouveau estimateur des moments modifiés notés MMe, nous allons utiliser l'espérance et poser

$$r = \left[\frac{1}{n} \sum_{i=1}^n X_i^{-1} \right]^{-1} \quad (2.5)$$

au lieu d'utiliser l'espérance et la variance. Nous aurons les deux équations :

$$\mu = \frac{\alpha}{\beta} \quad (2.6)$$

$$r^{-1} = \frac{\beta}{\alpha - 1}, \quad \alpha > 1 \quad (2.7)$$

Après résolu les équations pour α et β , nous obtenons le MMe pour α et β , notés $\tilde{\alpha}$ et $\tilde{\beta}$ respectivement.

$$\tilde{\alpha} = \frac{\mu}{\mu - r} = 1 + \frac{1}{\frac{\mu}{r-1}} \quad (2.8)$$

$$\tilde{\beta} = \frac{1}{\mu - r'} \quad (2.9)$$

Avec $\frac{\mu}{r} = \frac{1}{n^2} \sum_{i=1}^n \sum_{j=1}^n \frac{X_i}{X_j}$

Remarque 1. Pour les MMe $\tilde{\alpha}$ et $\tilde{\beta}$, α doit être supérieur à 1 puisque $E(X^{-1}) = \frac{\beta}{\alpha-1}$, $\alpha \geq 1$ La distribution jointe asymptotique de $\sqrt{n}(\tilde{\alpha} - \alpha, \tilde{\beta} - \beta)$ est $\mathcal{N}(0, \Sigma)$ avec

$$\Sigma = \begin{bmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{bmatrix}$$

Où

$$\Sigma_{11} = \frac{\beta^4}{\alpha^3(\alpha - 1)^2\Theta} \left(3 + \frac{\alpha}{\alpha - 2} \right)$$

$$\Sigma_{12} = \Sigma_{21} = \frac{\beta}{\alpha(\alpha-1)\Delta\Theta} \left[2 + \frac{\beta^2}{\alpha(\alpha-1)} + \frac{\beta^2}{(\alpha-2)(\alpha-1)} \right]$$

$$\Sigma_{22} = \frac{1}{\Delta^2} \left[\frac{\alpha}{\beta^2} + \frac{1}{\alpha-1} + \frac{\beta^2}{(\alpha-2)(\alpha-1)^2} \right]$$

Avec $\Theta = \left[1 - \frac{\beta^2}{\alpha(\alpha-1)} \right]^2$ et $\Delta = \left(\frac{\alpha}{\beta} - \frac{\beta}{\alpha-1} \right)^2$ Ce résultat peut-être prouvé en utilisant les étapes données dans l'appendix de NG, Kundu et Balakrishan(2003,p296-298).

Propriété 2. La méthode des moment modifiés donne des estimateurs consistants.

Preuve 1. μ et r sont des estimateurs consistants de $E(X) = \frac{\alpha}{\beta}$ et $E(X^{-1}) = (\alpha-1)/\beta$ (en utilisant la loi des grands nombres), respectivement, alors : $p \lim_{n \rightarrow \infty} \mu = \frac{\alpha}{\beta}$ et $p \lim_{n \rightarrow \infty} r = \frac{\alpha-1}{\beta}$ En prenant les limites en probabilité, nous obtenons :

$$p \lim_{n \rightarrow \infty} \tilde{\alpha} = p \lim_{n \rightarrow \infty} \frac{\mu}{\mu - r} = \frac{\alpha/\beta}{1/\beta} = \alpha \quad (2.10)$$

$$p \lim_{n \rightarrow \infty} \tilde{\beta} = p \lim_{n \rightarrow \infty} \frac{1}{\mu - r} = \frac{1}{1/\beta} = \beta \quad (2.11)$$

Conclusion 1. Nous concluons que $\tilde{\alpha}$ converge vers α , et $\tilde{\beta}$ converge vers β presque sûrement.

Nouvel estimateur des moments modifiés NMMe

Avec X_1, \dots, X_n un échantillon complet de taille n où X_i suit une loi $\Gamma(\alpha, \beta)$. Définissons la variable aléatoire $Z_{ij} = X_i X_j^{-1}$ pour $1 \leq i \neq j \leq n$, où $Z_{ij} = \frac{1}{Z_{ji}}$. Maintenant, il est possible de montrer que :

$$E(Z_{ij}) = E(X_i X_j^{-1}) = \frac{\alpha}{\beta} \frac{\beta}{\alpha-1} = \frac{\alpha}{\alpha-1}, \alpha > 1 \quad (2.12)$$

Aussi notons que la moyenne de l'échantillon est donnée par :

$$\bar{z} = \frac{1}{2n^2} \sum_{1 \leq i \neq j \leq n} Z_{ij} \quad (2.13)$$

Si nous mettons (2.12) = (2.13), nous obtenons

$$\tilde{\alpha}^* = \frac{\bar{z}}{\bar{z} - 1} = 1 + \frac{1}{(\mu/r)^* - 1} \quad (2.14)$$

Où $(\mu/r)^* = \frac{1}{n(n-1)} \sum_{1 \leq i \neq j \leq n} X_i X_j^{-1}$

Remarque 2. Les nouveaux estimateurs $\tilde{\alpha}^*$ et $\tilde{\beta}^*$ exigent que $\alpha > 1$ puisque $E(X^{-1}) = \frac{\beta}{\alpha-1}$ pour $\alpha > 1$.

Propriété 3. L'estimateur $\tilde{\alpha}^*$ est consistant.

Preuve 2. En prenant la limite en probabilité de l'équation (3.16), nous obtenons :

$$p \lim_{n \rightarrow \infty} \tilde{\alpha}^* = p \lim_{n \rightarrow \infty} \frac{\bar{z}}{\bar{z} - 1} = \frac{\frac{\alpha}{\alpha-1}}{\frac{1}{\alpha-1}} = \alpha \quad (2.15)$$

Propriété 4. $\frac{1}{\tilde{\alpha}^* - 1}$ est un estimateur sans biais de $\frac{1}{\alpha - 1}$.

Preuve 3. Notons que :

$$\frac{1}{\tilde{\alpha}^* - 1} = \frac{1}{1 + \left[\left(\frac{\mu}{r}\right)^* - 1\right]^{-1} - 1} = \left(\frac{\mu}{r}\right)^* - 1 \quad (2.16)$$

Ce qui implique que :

$$E\left(\frac{1}{\tilde{\alpha}^* - 1}\right) = \frac{\alpha}{\alpha - 1} - 1 = \frac{1}{\alpha - 1}$$

Propriété 5. $\tilde{\alpha}^*$ est un estimateur positivement biaisé de α

Preuve 4. Soit $f(x) = \frac{1}{x-1}$, alors $f'(x) = -\frac{1}{x-1}^2$ et $f''(x) = \frac{2}{x-1}^3 > 0$ pour $x > 1$. Alors, en utilisant l'inégalité de Jensen, nous avons :

$$f[E(\tilde{\alpha}^*)] \leq E[f(\tilde{\alpha}^*)] \Rightarrow \frac{1}{E(\tilde{\alpha}^*) - 1} \leq \frac{1}{\alpha - 1} \Rightarrow E(\tilde{\alpha}^*) - 1 \geq \alpha - 1 \Rightarrow E(\tilde{\alpha}^*) \geq \alpha \quad (2.17)$$

Propriété 6. Soit X_1, \dots, X_n un échantillon aléatoire de taille n , suivant une loi gamma (1.2), alors $\tilde{\alpha}^* < \tilde{\alpha}$

Preuve 5. D'abord nous avons $\bar{z} \geq \mu/r$ et donc :

$$\bar{z} - 1 \geq \mu/r - 1 = \frac{\mu - r}{r} \Rightarrow \frac{1}{\bar{z} - 1} \leq \frac{r}{\mu - r} = \frac{\mu}{\mu - r} - 1 \Rightarrow \frac{1}{\bar{z} - 1} + 1 \leq \frac{\mu}{\mu - r}$$

Alors,

$$\frac{\bar{z}}{\bar{z} - 1} \leq \frac{\mu}{\mu - r} \quad (2.18)$$

(voire [30] pour plus de détails)

Propriété 7. $\tilde{\alpha}^*$ aura toujours un biais et une variance plus petits que ceux de $\tilde{\alpha}$.

Preuve 6. La preuve vient directement des propriétés de la loi Gamma. Comme :

$$E(\mu) = E\left(\frac{1}{n} \sum_{i=1}^n X_i\right) = \frac{\alpha}{\beta}$$

Il est possible d'obtenir un estimateur de β noté $\tilde{\beta}^*$

$$\tilde{\beta}^* = \frac{\tilde{\alpha}^*}{\mu} = \frac{\bar{z}}{\mu(\bar{z} - 1)} \quad (2.19)$$

Propriété 8. L'estimateur $\tilde{\beta}^*$ est consistant.

Preuve 7. En utilisant la limite en probabilité de l'équation 2.19, nous obtenons :

$$p \lim_{n \rightarrow \infty} \tilde{\beta}^* = p \lim_{n \rightarrow \infty} \frac{\tilde{\alpha}^*}{\alpha/\beta} = \beta \quad (2.20)$$

Estimateur à biais réduit de α

Ici, nous dérivons la distribution asymptotique et le biais de l'estimateur $\tilde{\alpha}^*$. Pour la quantité $\left(\frac{\mu}{r}\right)^*$, nous trouvons $E\left[\left(\frac{\mu}{r}\right)^*\right] = E\left(\frac{X_i}{X_j}\right) = \frac{\alpha}{\alpha-1}$ et

$$E\left[\left(\frac{\mu}{r}\right)^{*2}\right] = \frac{1}{n^2(n-1)^2} E\left[\sum_{1 \leq i \neq j \neq k \neq l \leq n} \frac{X_j X_i}{X_k X_l} + \sum_{1 \leq i \neq j \neq k \leq n} \frac{X_i^2}{X_j X_k} + \sum_{1 \leq i \neq j \neq k \leq n} \frac{X_j X_k}{X_i^2} + 2 \sum_{1 \leq i \neq j \neq k \leq n} \frac{X_j X_i}{X_i X_k} + \sum_{1 \leq i \neq j \neq k \leq n} \frac{X_i^2}{X_j^2} + \sum_{1 \leq i \neq j \neq k \leq n} \frac{X_i X_j}{X_j X_i}\right]$$

$$= \frac{1}{n^2(n-1)^2} \left[\frac{\alpha^2}{\alpha^2 - 1} n(n-1)(n-2)(n-3) + n(n-1)(n-2) \left(\frac{\alpha(\alpha+1)}{(\alpha-1)^2} + \frac{\alpha^2}{(\alpha-1)(\alpha-2)} + \frac{2\alpha}{\alpha-1} \right) + n(n-1) \left(\frac{\alpha(\alpha+1)}{(\alpha-1)(\alpha-2)} + 1 \right) \right]$$

$$= \frac{(n-2)(n-3)\alpha^2}{n(n-1)(\alpha-1)^2} + \frac{(n-2)}{n(n-1)} \frac{2\alpha(2\alpha^2 - 4\alpha + 1)}{n(n-1)(\alpha-1)^2(\alpha-2)} + \frac{1}{n(n-1)} \frac{2(\alpha^2 - \alpha + 1)}{(\alpha-1)(\alpha-2)}$$

Ce résultat nous donne :

$$\begin{aligned} Var \left[\left(\frac{\mu}{r} \right)^* \right] &= \frac{(n-2)(n-3)\alpha^2}{n(n-1)(\alpha-1)^2} + \frac{(n-2)}{n(n-1)} \frac{2\alpha(2\alpha^2-4\alpha+1)}{(\alpha-1)^2(\alpha-2)} \\ &+ \frac{1}{n(n-1)} \frac{2(\alpha^2-\alpha+1)}{(\alpha-1)(\alpha-2)} - \frac{\alpha^2}{(\alpha-1)^2} \end{aligned}$$

Ce qui implique,

$$Var \left[\left(\frac{\mu}{r} \right)^* \right] \approx \frac{2\alpha}{(n-1)(\alpha-1)^2(\alpha-2)}$$

Pour obtenir la distribution de $\tilde{\alpha}^*$, nous pouvons utiliser l'expansion de Taylor avec $\bar{\omega} = E \left[\left(\frac{\mu}{r} \right)^* \right] = \frac{\alpha}{\alpha-1}$

$$\begin{aligned} \tilde{\alpha}^* &= 1 + \frac{1}{\left(\frac{\mu}{r} \right)^* - 1} = g\left(\left(\frac{\mu}{r}\right)^*\right) = g(\bar{\omega}) + \left(\left(\frac{\mu}{r} \right)^* - \bar{\omega} \right) g'(\bar{\omega}) + \frac{\left(\left(\frac{\mu}{r} \right)^* - \bar{\omega} \right)^2}{2} g''(\bar{\omega}) + \dots \\ &\approx 1 + (\alpha-1) - (\alpha-1)^2 \left[\frac{\mu^*}{r} - \frac{\alpha}{\alpha-1} \right] + (\alpha-1)^3 \left[\frac{\mu^*}{r} - \frac{\alpha}{\alpha-1} \right]^2 \end{aligned} \quad (2.21)$$

Où $g'(\cdot)$ et $g''(\cdot)$ sont la première et seconde dérivée respectivement de la fonction $g(\cdot)$. Par conséquent, en utilisant la méthode Delta, la distribution asymptotique de $\tilde{\alpha}^*$ est $\mathcal{N} \left(\alpha, \frac{2\alpha(\alpha-1)^2}{(n-1)(\alpha-2)} \right)$

Ainsi de (2.21), nous avons

$$Biais(\tilde{\alpha}^*) \approx \frac{2\alpha(\alpha-1)}{(n-1)(\alpha-2)} \quad (2.22)$$

Ce qui nous permet de proposer un estimateur de biais réduit pour α (pour $\alpha > 2$), noté par $\tilde{\alpha}^{**}$, comme :

$$\tilde{\alpha}^{**} = \tilde{\alpha}^* - \tilde{Biais}(\tilde{\alpha}^*)$$

Où le $\tilde{Biais}(\tilde{\alpha}^*)$ est une estimation du biais $\tilde{\alpha}^*$, i.e. le paramètre inconnu dans le biais est remplacé par son NMME correspondant.

2.1.2 Méthode du Maximum de vraisemblance

La méthode du maximum de vraisemblance a été proposée par *R.A Fisher* en 1912 puis il l'a développée en 1920. L'estimateur de maximum de vraisemblance de la loi Gamma quant à lui, a été discuté par beaucoup d'auteurs tels que Choi et Wette 1969, Gross et Shenton, Hwang et Huang 2022.

Définition 2. Soit un n -échantillon (X_1, X_2, \dots, X_n) issu d'une variable aléatoire X de loi de probabilité $f(x, \theta), \theta \in \Theta$. On appelle fonction de vraisemblance de θ pour une réalisation de l'échantillon (x_1, x_2, \dots, x_n) la fonction :

$$\mathcal{L}(\theta, x_1, x_2, \dots, x_n) = \prod_{i=1}^n f(x_i, \theta)$$

Définition 3. La statistique :

$$T \longrightarrow \operatorname{argmax}_{\theta \in \Theta} \left(\prod_{i=1}^n f(x_i, \theta) \right)$$

s'appelle **estimateur du maximum de vraisemblance** noté $\hat{\theta}_{MV}$.

Souvent, la fonction de vraisemblance a une écriture complexe, le logarithme lui est appliqué pour la simplifier, on obtient ainsi la fonction dite *log-vraisemblance* définie par :

$$l = l(\theta, x_1, x_2, \dots, x_n) = \log \mathcal{L}(\theta, x_1, x_2, \dots, x_n)$$

L'estimateur du MV est obtenu en résolvant le système suivant :

$$\begin{cases} \frac{\partial f}{\partial \theta} = 0 \\ \frac{\partial^2 f}{\partial \theta^2} < 0 \end{cases}$$

Soit (X_1, \dots, X_n) des variables aléatoires suivant une loi Gamma de paramètres α et β ,

La fonction de vraisemblance de la loi Gamma

$$\mathcal{L}(\alpha, \beta; x_1, x_2, \dots, x_n) = \prod_{i=1}^n f(x_i; \alpha, \beta)$$

$$\mathcal{L}(\alpha, \beta; x_1, x_2, \dots, x_n) = \frac{\beta^{n\alpha}}{(\Gamma(\alpha))^n} \prod_{i=1}^n x_i^{\alpha-1} e^{-\beta \sum_{i=1}^n x_i}$$

La log-vraisemblance

$$\log \mathcal{L} = \log \prod_{i=1}^n f(x_i; \alpha, \beta)$$

$$\log \mathcal{L} = \log \frac{\beta^{n\alpha}}{(\Gamma(\alpha))^n} \prod_{i=1}^n x_i^{\alpha-1} e^{-\beta \sum_{i=1}^n x_i}$$

$$\log \mathcal{L} = -n \log \Gamma(\alpha) + n\alpha \log \beta + (\alpha - 1) \sum_{i=1}^n \log x_i - \beta \sum_{i=1}^n x_i$$

Les paramètres maximisant la fonction de vraisemblance sont les solutions des équations suivantes :

$$\begin{cases} \frac{\partial \log \mathcal{L}}{\partial \alpha} = -n\Psi(\alpha) + n \log \beta + \sum_{i=1}^n \log x_i \\ \frac{\partial \log \mathcal{L}}{\partial \beta} = \frac{n\alpha}{\beta} - \sum_{i=1}^n x_i \end{cases}$$

Avec la fonction $\Psi = \frac{\partial^2}{\partial x^2} \log \Gamma(x)$ où Ψ' est la fonction tri-gamma.

Nous constatons avec les équations précédentes qu'elles sont très complexes à résoudre, ce qui rend l'obtention du maximum de vraisemblance difficile à trouver analytiquement pour α et β . Pour remédier à ça nous allons utiliser une méthode numérique pour obtenir des estimations pour α et β , qui seront respectivement $\hat{\alpha}$ et $\hat{\beta}$, qui vont maximiser la fonction de vraisemblance. Nous choisissons ici la méthode de *Newton-Raphson* en utilisant la matrice Hessienne (qui est la deuxième dérivée partielle de la log-vraisemblance).

On construit la matrice Hessienne comme suit :

Posant $g_1(\alpha)$ et $g_2(\beta)$ tels que :

$$g_1(\alpha) = -n\Psi(\alpha) + n \log \beta + \sum_{i=1}^n \log x_i$$

$$g_2(\beta) = \frac{n\alpha}{\beta} - n\bar{X}$$

Les dérivées partielles respectives sont :

– pour g_1

$$\frac{\partial g_1(\alpha)}{\partial(\alpha)} = -n\Psi'(\alpha)$$

$$\frac{\partial g_1(\alpha)}{\partial(\beta)} = \frac{n}{\beta}$$

— pour g_2

$$\frac{\partial g_2(\beta)}{\partial(\alpha)} = \frac{n}{\beta}$$

$$\frac{\partial g_2(\beta)}{\partial(\beta)} = \frac{-n\alpha}{\beta^2}$$

La Hessienne nous donnera donc :

$$J_k = \begin{bmatrix} \frac{\partial g_1(\alpha)}{\partial \alpha} & \frac{\partial g_1(\alpha)}{\partial \beta} \\ \frac{\partial g_2(\beta)}{\partial \alpha} & \frac{\partial g_2(\beta)}{\partial \beta} \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$$

Où J_k est la matrice Jacobienne, qui doit être inversible et symétrique.

L'inverse J_k noté J_k^{-1} est :

$$J_k^{-1} = \frac{1}{|J|} \begin{bmatrix} a_{11} & -a_{12} \\ -a_{21} & a_{22} \end{bmatrix}$$

$$\begin{bmatrix} \alpha_{k+1} \\ \beta_{k+1} \end{bmatrix} = \begin{bmatrix} \alpha_k \\ \beta_k \end{bmatrix} - J_k^{-1} \begin{bmatrix} g_1(\alpha) \\ g_2(\beta) \end{bmatrix}$$

$$= \begin{bmatrix} \alpha_k \\ \beta_k \end{bmatrix} - \frac{1}{a_{11}a_{22} - a_{12}a_{21}} \begin{bmatrix} -na_{22}\Psi(\alpha_k) + na_{22} \log \beta_k + a_{22} \sum_{i=1}^n \log x_i - \frac{na_{12}\alpha_k}{\beta_k} + na_{12}\bar{x} \\ na_{21}\Psi(\alpha_k) - na_{21} \log \beta_k - a_{21} \sum_{i=1}^n \log x_i + \frac{na_{11}\alpha_k}{\beta_k} + na_{11}\bar{x} \end{bmatrix}$$

L'écart entre les valeurs de la première itération avec la dernière désigne l'erreur résiduelle, notée ϵ , qui est une valeur minime donnée par :

$$\begin{bmatrix} \epsilon_{k+1}(\alpha) \\ \epsilon_{k+1}(\beta) \end{bmatrix}$$

Qui est égale à :

$$= \left| \begin{bmatrix} \alpha_{k+1} \\ \beta_{k+1} \end{bmatrix} - \begin{bmatrix} \alpha_k \\ \beta_k \end{bmatrix} \right|$$

Où α_k et β_k sont les valeurs initiales pour α et β respectivement.

Propriété 9. *L'estimateur de maximum de vraisemblance est biaisé dans les échantillons finis.*

2.2 Estimation Bayésienne

L'inférence bayésienne est une approche statistique qui permet de tirer des conclusions à partir des données en utilisant des probabilités. Elle repose sur le théorème de Bayes, qui permet de mettre à jour nos croyances initiales à la lumière de nouvelles données. Dans cette inférence, les paramètres inconnus sont traités comme des variables aléatoires et sont décrits par des distributions de probabilité appelées distributions a priori. Ces distributions a priori expriment nos connaissances, nos croyances ou nos incertitudes avant d'observer les données.

2.2.1 Modèle Bayésien

Définition 4. On appelle modèle statistique Bayésien la donnée d'un modèle statistique paramétré $(\mathcal{X}, \mathcal{A}, \mathcal{P}_\theta)$, où $\theta \in \Theta$.

Définissons les espaces intervenant dans un modèle Bayésien :

Espace des observations Noté \mathcal{X} , il représente l'ensemble des résultats suite à une étude d'un phénomène.

Espace des actions Noté \mathcal{A} , il représente l'ensemble des actions ou décisions à prendre après l'obtention de l'information.

Espace des états de la nature Noté Θ , il représente l'espace des paramètres inconnus θ .

Espace des règles de décisions Noté D , il représente l'ensemble des règles de décisions qu'on définit comme une application de \mathcal{X} dans \mathcal{A} notée par δ .

$$\delta : \mathcal{X} \longrightarrow \mathcal{A}$$

$$x_i \longrightarrow \delta(x_i) = a_i$$

Les lois qui interviennent dans la statistique bayésienne :

Définition 5. Loi a priori

Notée par $\pi(\theta)$, désigne la distribution du paramètre inconnu θ , elle porte l'information sur ce dernier. L'appellation a priori exprime le fait qu'elle a été établie préalablement à l'observation des données

Définition 6. Loi du couple

Cette loi est appelée loi jointe, sa densité est notée $h(\theta, x)$

on a

$$h(\theta, x) = f(X|\theta) \cdot \pi(\theta)$$

Définition 7. Loi marginale de X

La loi marginale de X notée $f(x)$ est calculée de la manière suivante :

$$f(x) = \int_{\Theta} h(\theta, x) d\theta$$

Définition 8. Loi a posteriori

C'est la loi constitutionnelle de θ sachant X, sa densité est noté $\pi(\theta|X)$, cette loi est une actualisation de la loi a priori $\pi(\theta)$ aux vu de l'observation x_i . En vertu du théorème de Bayes, on a

$$\pi(\theta|X) = \frac{f(X|\theta) \cdot \pi(\theta)}{f(X)}$$

où

$f(X|\theta)$ est la fonction de vraisemblance,

Et $f(X) = \int f(X|\theta) \cdot \pi(\theta) d\theta$.

2.2.2 Théorème de Bayes

Définition 9. Théorème de Bayes

Considérons les événements aléatoires A et B tels que $P[B] \neq 0$. $P(A|B)$ et $P(B|A)$ sont définis respectivement comme étant la probabilité de A , conditionnellement à la réalisation de B , et la probabilité de B , conditionnellement à la réalisation de A , de la manière suivante :

$$P[A|B] = \frac{P[A, B]}{P[B]}$$

et

$$P[B|A] = \frac{P[B, A]}{P[A]}$$

En utilisant la relation $P(A, B) = P(B, A)$, qui représente la probabilité que les événements A et B se produisent simultanément, on peut établir la relation entre les deux probabilités conditionnelles $P(A|B)$ et $P(B|A)$.

En remplaçant $P(A|B)$ par $P(B|A)P(A)$, qui est équivalent, on peut déduire la relation entre les deux probabilités conditionnelles $P(A|B)$ et $P(B|A)$:

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)}$$

Cette relation, connue sous le nom de théorème de Bayes, représente en réalité un principe d'actualisation. Elle décrit comment la vraisemblance de $P(A)$ est mise à jour vers $P(A|B)$ une fois que l'événement B a été observé.

Bayes (1763) a formulé une version continue de ce résultat : pour deux variables aléatoires X et Y avec une distribution conditionnelle $f(x|y)$ et une distribution marginale $g(y)$, la distribution conditionnelle de y sachant x est donnée par :

$$g(y|x) = \frac{f(x|y)g(y)}{\int f(x|y)g(y)dy}$$

Ce théorème d'inversion est naturellement justifié d'un point de vue probabiliste. Cependant, Bayes et Laplace sont allés plus loin en considérant que l'incertitude concernant le paramètre θ d'un modèle peut être décrit par une distribution de

probabilité π sur Θ , appelée distribution a priori. L'inférence est alors basée sur la distribution conditionnelle de θ étant donné x , $\pi(\theta|x)$, qui est appelée distribution a posteriori et est définie par :

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

La formule de Bayes, également connue sous le nom d'équation de Bayes, peut être présentée de la manière suivante, étant donné que le dénominateur est indépendant de θ :

$$\pi(\theta|x) \propto f(x|\theta)\pi(\theta)$$

La transition de la distribution a priori à la distribution a posteriori des paramètres du modèle statistique, exprimée par la formule de Bayes, peut être interprétée comme une mise à jour ou une actualisation des connaissances en fonction des observations.

2.2.3 Fonctions coût usuelles

L'objectif des études inférentielles est de fournir une décision au statisticiens.

Les différentes décisions sont comparées au moyen d'un critère d'évolution donné par la fonction coût .

Définition 10. *On appelle fonction coût toute fonction L :*

$$\begin{cases} \Theta \times \mathcal{A} \longrightarrow \mathbb{R}^+ \\ ((\theta, \delta(x)) \longrightarrow L(\theta, \delta(x))) \end{cases}$$

$L(\theta, \delta(x))$ évalue le coût ou la perte associé a la décision $a = \delta(x)$ quand le paramètre vaut θ . Elle permet de calculer la perte pour une mauvaise décision.

Nous allons définir 3 fonctions coûts usuelles, elles sont appréciés pour leurs utilité et simplicités, elles nous permettent d'obtenir aisément l'estimateur Bayésien.

Le coût quadratique

Définition 11. La fonction coût quadratique est définie par :

$$L(\theta, \delta) = (\theta - \delta)^2$$

une variante de cette fonction coût est le coût quadratique pondérée définie par :

$$L(\theta, \delta) = g(\theta)(\theta - \delta)^2.$$

Proposition 1. L'estimateur de Bayes δ^π associé à la loi a priori π et à la fonction coût quadratique est la moyenne a posteriori de θ . Cet estimateur est appelé Maximum Mean Squared Error (MMSE) donc :

$$\delta^\pi(x) = E[\theta|X].$$

La fonction coût absolue

Définition 12. La fonction coût absolue est définie par :

$$L(\theta, \delta) = \begin{cases} \theta - \delta & \text{si } \theta > \delta \\ \delta - \theta & \text{si } \theta \leq \delta \end{cases} \Rightarrow L(\theta, \delta) = |\theta - \delta|.$$

Proposition 2. L'estimateur de Bayes δ^π associé à la loi a priori π et à la fonction coût absolue est la médiane de $\pi(\theta|X)$.

La fonction coût 0-1

Définition 13. La fonction coût 0-1 est définie par :

$$L(\theta, \delta) = \begin{cases} 0 & \text{si } |\theta - \delta| < \varepsilon \\ 1 & \text{si } |\theta - \delta| \geq \varepsilon \end{cases} \quad \forall \varepsilon > 0.$$

2.2.4 Estimateur de Bayes

Définition 14. Nous appelons estimateur de Bayes associé à une fonction coût L et à une distribution a priori π , toute décision δ^π qui minimise le risque de Bayes $r(\pi, \delta)$.

on a :

$$\delta^\pi = r(\pi) = r(\pi, \delta^\pi) = \inf_{\delta \in D} r(\pi, \delta)$$

Le tableau ci-dessous représente quelques estimateurs de Bayes du paramètre θ sous coût quadratique pour les lois a priori conjuguées des familles exponentielles usuelles.

Loi de x	Loi conjuguée	Moyenne a posteriori
Normale $\mathcal{N}(\theta, \sigma^2)$	Normale $\mathcal{N}(\mu, \tau^2)$	$\frac{(\sigma^2\mu + \tau^2x)}{\sigma^2 + \tau^2}$
Poisson $\mathcal{P}(\theta)$	Gamma $\Gamma(\alpha, \beta)$	$\frac{\alpha + x}{\beta + 1}$
Gamma $\Gamma(v, \theta)$	Gamma $\Gamma(\alpha, \beta)$	$\frac{\alpha + v}{\beta + x}$
Binomiale $\mathcal{B}(n, \theta)$	Bêta $\mathcal{B}e(\alpha, \beta)$	$\frac{\alpha + x}{\alpha + \beta + n}$
Binomiale Négative Neg (n, θ)	Bêta $\mathcal{B}e(\alpha, \beta)$	$\frac{\alpha + n}{\alpha + \beta + x + n}$
Multinomiale $\mathcal{M}_k(n; \theta_1, \dots, \theta_k)$	Dirichlet $\mathcal{D}(\alpha_1, \dots, \alpha_k)$	$\frac{\alpha_i + x_i}{(\sum_j \alpha_j + n)}$
Normale $\mathcal{N}(\mu, \frac{1}{\theta})$	Gamma $\Gamma(\alpha/2, \beta/2)$	$\frac{\alpha + 1}{\beta + (\mu - x)^2}$

TABLE 2.1 – Quelques estimateurs de Bayes usuels.

Définition 15. (Risque de Bayes)

Le risque Bayésien noté $r(\pi, \delta)$, est défini pour une fonction de perte donnée par :

$$\begin{aligned} r(\pi, \delta) &= E^\pi[R(\theta, \delta)] \\ &= \int_{\theta} R(\theta, \delta) \pi(\theta) d\theta \end{aligned}$$

Ainsi, le risque de Bayes $r(\pi, \delta)$ peut également être exprimé comme la moyenne du risque a posteriori $\rho(\theta, \delta(x))$ suivant la loi marginale $f(x)$.

2.3 Estimation Bayésienne des paramètres de la loi Gamma

En utilisant les lois a priori Gamma (alpha,beta) et Exponentielle de paramètres c pour estimer les paramètres α et β d'une loi Gamma, nous supposons que α a une loi a priori $\pi_1(\alpha)$ qui suit une loi Gamma(a,b). On ne suppose aucune loi a priori spécifique sur α , on suppose simplement que la loi a priori de β est $\pi_2(\cdot)$ et la fonction de densité de $\pi_2(\cdot)$ est exponentielle, indépendante de $\pi_1(\cdot)$.

$$\pi_1(\alpha) = \begin{cases} \frac{(b)^a (\alpha)^{a-1} e^{-b\alpha}}{\Gamma(a)} & \text{si } a > 0, b > 0, \alpha > 0 \\ 0 & \text{sinon} \end{cases} \quad (2.23)$$

$$\pi_2(\beta) = \begin{cases} ce^{-\beta c} & \text{si } c > 0, \beta > 0 \\ 0 & \text{sinon} \end{cases} \quad (2.24)$$

Les équations précédentes sont les lois a priori de α et β respectivement.

La fonction de densité conjointe est donnée par la formule suivante :

$$\begin{aligned} \mathcal{J}(x_1, x_2, \dots, x_n; \alpha, \beta) &= \mathcal{L}(x_1, x_2, \dots, x_n; \alpha, \beta) \pi_1(\alpha) \pi_2(\beta) \\ &= \frac{\beta^{n\alpha}}{\Gamma(\alpha)^n} \prod_{i=1}^n x_i^{\alpha-1} e^{-\beta \sum_{i=1}^n x_i} \frac{(b)^a (\alpha)^{a-1} e^{-b\alpha}}{\Gamma(a)} ce^{-\beta c} \end{aligned} \quad (2.25)$$

La fonction de densité marginale est

$$f(x_1, x_2, \dots, x_n) = \int_0^\infty \int_0^\infty \mathcal{L}(x_1, x_2, \dots, x_n; \alpha, \beta) \pi_1(\alpha) \pi_2(\beta) d\alpha d\beta \quad (2.26)$$

Par conséquent :

la fonction de densité a posteriori de α et β peut-être écrite comme suit :

$$\begin{aligned} h(\alpha, \beta | x_1, x_2, \dots, x_n) &= \frac{\mathcal{L}(x_1, x_2, \dots, x_n; \alpha, \beta) \pi_1(\alpha) \pi_2(\beta)}{\int_0^\infty \int_0^\infty \mathcal{L}(x_1, x_2, \dots, x_n; \alpha, \beta) \pi_1(\alpha) \pi_2(\beta) d\alpha d\beta} \\ &= \frac{\frac{\beta^{n\alpha}}{\Gamma(\alpha)^n} \prod_{i=1}^n x_i^{\alpha-1} e^{-\beta \sum_{i=1}^n x_i} \frac{(b)^a (\alpha)^{a-1} e^{-b\alpha}}{\Gamma(a)} ce^{-\beta c}}{\int_0^\infty \int_0^\infty \frac{\beta^{n\alpha}}{\Gamma(\alpha)^n} \prod_{i=1}^n x_i^{\alpha-1} e^{-\beta \sum_{i=1}^n x_i} \frac{(b)^a (\alpha)^{a-1} e^{-b\alpha}}{\Gamma(a)} ce^{-\beta c} d\alpha d\beta} \end{aligned} \quad (2.27)$$

2.3.1 Les estimateurs Bayésiens sous différentes fonctions coût

Pour l'estimation Bayésienne des paramètres α et β de la loi Gamma, nous allons utiliser trois types de fonctions,

Fonction coût "précaution"

La fonction coût "précaution", a été introduite par Norstrom en 1996, cette fonction est asymétrique et peut-être définie par :

$$L_1 = \frac{(\theta - \hat{\theta})^2}{\hat{\theta}} \quad (2.28)$$

Basé sur cette fonction coût, le risque Bayésien est calculé comme suit :

$$\begin{aligned} R_B(\theta; \hat{\theta})E[L(\theta; \hat{\theta})] &= \int_0^\infty L(\theta; \hat{\theta})h(\theta|\underline{x})d\theta \\ &= \int_0^\infty \frac{(\theta - \hat{\theta})^2}{\hat{\theta}}h(\theta|\underline{x})d\theta \\ &= \int_0^\infty (\hat{\theta} - \hat{\theta}^{-1})h(\theta|\underline{x})d\theta - \int_0^\infty 2\theta h(\theta|\underline{x})d\theta + \int_0^\infty \hat{\theta}h(\theta|\underline{x})d\theta \\ R_B(\theta; \hat{\theta}) &= E(\theta^2|\underline{x})\hat{\theta}^{-1} - 2E(\theta|\underline{x}) + \hat{\theta} \end{aligned} \quad (2.29)$$

En prenant la dérivée partielle pour $R_B(\theta; \hat{\theta})$ et la mettant égale à 0, cela nous donne $\hat{\theta} = E(\theta^2|\underline{x})$

Par conséquent, l'estimateur de Bayes relatif à la fonction coût "précaution" notée $\hat{\theta}_B$ est donné par

$$\hat{\theta}_1 = \sqrt{E(\theta^2|\underline{x})} \quad (2.30)$$

En général :

$$E[(u(\alpha, \beta))] = \int_0^\infty \int_0^\infty u(\alpha, \beta)h(\alpha, \beta|x_1, \dots, x_n)d\alpha d\beta \quad (2.31)$$

Où $u(\alpha, \beta)$ peut-être n'importe quelle fonction de α et β . Donc :

$$E[(u(\alpha, \beta))] = \frac{\int_0^\infty \int_0^\infty u(\alpha, \beta)\mathcal{L}(x_1, x_2, \dots, x_n; \alpha, \beta)\pi_1(\alpha)\pi_2(\beta)d\alpha d\beta}{\int_0^\infty \int_0^\infty \mathcal{L}(x_1, x_2, \dots, x_n; \alpha, \beta)\pi_1(\alpha)\pi_2(\beta)d\alpha d\beta} \quad (2.32)$$

Estimation bayésienne du paramètre α Pour estimer le paramètre α sous la fonction coût "précaution", nous supposons que $u(\alpha, \beta) = \alpha^2$, donc nous aurons :

$$E(\alpha^2|x) = \frac{\int_0^\infty \int_0^\infty \alpha^2 \mathcal{L}(x_1, x_2, \dots, x_n; \alpha, \beta) \pi_1(\alpha) \pi_2(\beta) d\alpha d\beta}{\int_0^\infty \int_0^\infty \mathcal{L}(x_1, x_2, \dots, x_n; \alpha, \beta) \pi_1(\alpha) \pi_2(\beta) d\alpha d\beta} \quad (2.33)$$

Nous remarquons qu'il est complexe de trouver le ratio de deux intégrales double, alors nous allons utiliser l'approximation de *Lindley* pour trouver $E(\alpha^2|x)$.

Méthode de Lindley Nous avons $u(a, b) = \alpha^2$
Posons alors :

$$u_1 = \frac{\partial u(a,b)}{\partial \alpha} = 2\alpha, \quad u_{11} = \frac{\partial^2 u(a,b)}{\partial \alpha^2} = 2, \\ u_2 = \frac{\partial u(a,b)}{\partial \beta} = 0, \quad u_{22} = \frac{\partial^2 u(a,b)}{\partial \beta^2} = 0$$

et :

$$\pi(\alpha, \beta) = \frac{(b)^a (\alpha)^{a-1} e^{-b\alpha}}{\Gamma(a)} c e^{-\beta c}$$

Nous allons déduire alors :

$$P = \log \pi(\alpha, \beta) = (a-1) \log \alpha + a \log b - b\alpha - \log \Gamma(\alpha) + \log c - c\beta$$

$$P_1 = \frac{\partial P}{\partial \alpha} = \frac{a-1}{\alpha} \text{ et } P_2 = \frac{\partial P}{\partial \beta} = -c$$

Souvenons nous que

$$\log \mathcal{L}(x_1, x_2, \dots, x_n; \alpha, \beta) = n\alpha \log \beta - n \log \Gamma(\alpha) - \beta \sum_{i=1}^n x_i + (\alpha-1) \sum_{i=1}^n \log x_i$$

Alors :

$$\mathcal{L}_{12} = \frac{\partial^2 \log \mathcal{L}(\alpha, \beta)}{\partial \alpha \partial \beta^2} = \frac{-n}{\beta}, \quad l_{21} = \frac{\partial^3 \log l(\alpha, \beta)}{\partial^2 \partial \beta} = 0, \\ l_{03} = \frac{\partial^3 \log l(\alpha, \beta)}{\partial \beta^3} = \frac{2n\alpha}{\beta^3}, \quad l_{30} = \frac{\partial^3 \log l(\alpha, \beta)}{\partial \beta^3} = -n\Psi^n(\alpha) \\ \sigma_{11} = \frac{-1}{l_{20}} = \frac{1}{n\Psi'(\alpha)}, \quad \sigma_{22} = \frac{-1}{l_{02}} = \frac{\beta^2}{n\alpha}$$

En remplaçant dans $E(\alpha^2)$ nous aurons :

$$\begin{aligned}
E(\alpha^2) &\approx \hat{\alpha}^2 + \frac{1}{2} (u_1 \sigma_{11}) + p_1 u_{11} \sigma_{11} + \frac{1}{2} (l_{30} u_{11} \sigma_{11}^2) + \frac{1}{2} (l_{11} \sigma_{11} \sigma_{21}) \\
&\approx \hat{\alpha}^2 + \frac{1}{2} \left(2 \frac{1}{n \Psi'(\hat{\alpha})} \right) + \left(\frac{a-1}{\hat{\alpha}} - b \right) 2\hat{\alpha} \frac{1}{n \Psi'(\hat{\alpha})} \\
&\quad + \frac{1}{2} \left(-n \Psi''(\hat{\alpha}) \frac{2\hat{\alpha}}{(n \Psi'(\hat{\alpha}))^2} \right) + \frac{1}{2} \left(\frac{-n}{\beta^2} \frac{2\hat{\alpha}}{n \Psi'(\hat{\alpha})} \frac{\hat{\beta}^2}{n \hat{\alpha}} \right) \\
&\approx \hat{\alpha}^2 + \frac{2\hat{\alpha}}{n \Psi'(\hat{\alpha})} \left(\frac{a-1}{\hat{\alpha}} - b \right) - \frac{n \Psi''(\hat{\alpha}) \hat{\alpha}}{(n \Psi'(\hat{\alpha}))^2}
\end{aligned}$$

En remplaçant ce résultat de $E(\alpha^2)$ dans $\hat{\theta}_B$, nous allons obtenir l'estimateur de α :

$$\hat{\alpha}_1 = \sqrt{\hat{\alpha}^2 + \frac{2\hat{\alpha}}{n \Psi'(\hat{\alpha})} \left(\frac{a-1}{\hat{\alpha}} - b \right) - \frac{n \Psi''(\hat{\alpha}) \hat{\alpha}}{(n \Psi'(\hat{\alpha}))^2}} \quad (2.34)$$

Remarque Plus de détails et d'informations sont disponibles sur la méthode de Lindley dans l'annexe .1 à la fin du mémoire.

Estimation bayésienne du paramètre β Pour estimer le paramètre β , nous allons procéder de la même manière que pour le paramètre α ci-dessus.

Supposons que $u(\alpha, \beta) = \beta^2$, nous allons alors calculer $E(\beta^2)$:

$$\begin{aligned}
u_1 &= \frac{\partial u(a,b)}{\partial \alpha} = 0, & u_{11} &= \frac{\partial^2 u(a,b)}{\partial \alpha^2} = 0, \\
u_2 &= \frac{\partial u(a,b)}{\partial \beta} = 2\beta, & u_{22} &= \frac{\partial^2 u(a,b)}{\partial \beta^2} = 2
\end{aligned}$$

En remplaçant dans $E(\beta^2)$, nous aurons :

$$\begin{aligned}
E(\beta^2) &\approx \hat{\beta}^2 + \frac{1}{2} (u_{22} \sigma_{22}) + p_2 u_2 \sigma_{22} + \frac{1}{2} (l_{03} u_2 \sigma_{22}^2) \\
&\approx \hat{\beta}^2 + \frac{1}{2} \left(\frac{2\hat{\beta}^2}{n \hat{\alpha}} \right) + \left(-c \frac{2\hat{\beta}^3}{n \hat{\alpha}} + \frac{1}{2} \left(\frac{2n \hat{\alpha}}{\hat{\beta}^3} \frac{\hat{\beta}^4}{n \hat{\alpha}} \right) \right) \\
&\approx \hat{\beta}^2 + \frac{3\hat{\beta}^2}{n \hat{\alpha}} - \frac{2c\hat{\beta}^3}{n \hat{\alpha}}
\end{aligned}$$

En remplaçant $E(\beta^2)$ dans $\hat{\theta}_B$, nous aurons :

$$\hat{\beta}_1 \approx \sqrt{\hat{\beta}^2 + \frac{3\hat{\beta}^2}{n\hat{\alpha}} - \frac{2c\hat{\beta}^3}{n\hat{\alpha}}} \quad (2.35)$$

Où $\hat{\alpha}$ et $\hat{\beta}$ sont les estimateurs de maximum de vraisemblance de α et β respectivement.

Fonction coût d'entropie

La fonction coût d'entropie est définie par :

$$L_2(\hat{\theta}, \theta) = \left(\frac{\hat{\theta}}{\theta}\right) - \ln\left(\frac{\hat{\theta}}{\theta}\right) - 1 \quad (2.36)$$

L'estimateur de Bayes En utilisant la même méthode que celle utilisée avec la fonction coût, nous aurons :

$$\hat{\theta}_2 = [E(\theta^{-1}|X)]^{-1} \quad (2.37)$$

Estimateur de Bayes pour α Soit $u(\alpha, \beta)$ une fonction de α et β , alors $u(\alpha, \beta) = \frac{1}{\alpha}$, donc

$$u_1 = \frac{\partial u(\alpha, \beta)}{\partial \alpha} = -\alpha^{-2}$$

$$u_{11} = \frac{\partial^2 u(\alpha, \beta)}{\partial \alpha^2} = 2\alpha^{-3}$$

$$u_2 = \frac{\partial u(\alpha, \beta)}{\partial \beta} = 0$$

$$u_{22} = \frac{\partial^2 u(\alpha, \beta)}{\partial \beta^2} = 0$$

Alors

$$E\left(\frac{1}{\alpha}\right) \approx \frac{1}{\hat{\alpha}} + \frac{1}{2}(u_{11}\sigma_{11}) + p_1 u_1 \sigma_{11} + \frac{1}{2}(L_{30} u_1 \sigma_{11}^2) + \frac{1}{2}(L_{12} u_1 \sigma_{11} \sigma_{22})$$

$$E\left(\frac{1}{\alpha}\right) \approx \frac{1}{\hat{\alpha}} + \frac{1}{\hat{\alpha}^3 n \Phi'(\hat{\alpha})} - \frac{1}{\hat{\alpha}^2 n \Phi'(\hat{\alpha})} \left(\frac{a^{-1}}{\hat{\alpha}} - b\right) + \frac{1}{2} \left(\frac{n \Phi''(\hat{\alpha})}{\hat{\alpha}^2 (n \Phi'(\hat{\alpha}))^2}\right) + \frac{1}{2} \left(\frac{1}{\hat{\alpha}^3 n \Phi'(\hat{\alpha})}\right) \quad (2.38)$$

Maintenant, en remplaçant (2.38) dans (2.37) nous aurons :

$$\hat{\alpha}^2 \approx \frac{1}{\frac{1}{\hat{\alpha}} + \frac{1}{\hat{\alpha}^3 n \Phi'(\hat{\alpha})} - \frac{1}{\hat{\alpha}^2 n \Phi'(\hat{\alpha})} (a^{-1} \hat{\alpha} - b) + \frac{1}{2} \left(\frac{n \Phi''(\hat{\alpha})}{\hat{\alpha}^2 (n \Phi'(\hat{\alpha}))^2}\right) + \frac{1}{2} \left(\frac{1}{\hat{\alpha}^3 n \Phi'(\hat{\alpha})}\right)} \quad (2.39)$$

Estimateur de Bayes pour β En utilisant la même méthode, nous aurons

$u(\alpha, \beta) = \frac{1}{\beta}$, donc

$$u_1 = \frac{\partial u(\alpha, \beta)}{\partial \alpha} = 0$$

$$u_{11} = \frac{\partial^2 u(\alpha, \beta)}{\partial \alpha^2} = 0$$

$$u_2 = \frac{\partial u(\alpha, \beta)}{\partial \beta} = -\beta^{-2}$$

$$u_{22} = \frac{\partial^2 u(\alpha, \beta)}{\partial \beta^2} = 2\beta^{-3}$$

Alors

$$E\left(\frac{1}{\beta}\right) \approx \frac{1}{\hat{\beta}} + \frac{1}{2}(u_{22}\sigma_{22}) + p_2 u_2 \sigma_{22} + \frac{1}{2}(L_{03} u_2 \sigma_{22}^2) + \frac{1}{2}(L_{21} u_2 \sigma_{11} \sigma_{22}) \quad (2.40)$$

$$E\left(\frac{1}{\beta}\right) \approx \frac{1}{\hat{\beta}} + \frac{1}{2} \left(2\hat{\beta}^2 \frac{\hat{\beta}^3 n \hat{\alpha}}{\beta}\right) + c \hat{\beta}^2 \frac{\beta^2 n \hat{\alpha}}{2} + \frac{1}{2} \left(-2n \hat{\alpha} \frac{\hat{\beta}^3}{\beta^2} \frac{\beta^4}{n^2 \hat{\alpha}^2}\right) \approx \frac{1}{\hat{\beta}} + cn \hat{\alpha}() \quad (2.41)$$

En remplaçant (2.41) dans (2.37), nous obtenons

$$\hat{\beta}^2 \approx \frac{1}{\frac{1}{\hat{\beta}} + \frac{c}{n\hat{\alpha}}} \quad (2.42)$$

Où $\hat{\beta}$ et $\hat{\alpha}$ sont les estimateurs de maximum de vraisemblance

Fonction coût quadratique pondérée

Cette fonction coût a été discutée par DeGroot(1970). C'est une fonction symétrique, positive et continue. Elle est définie par :

$$L_3(\theta, \hat{\theta}) = \left(\frac{\theta - \hat{\theta}}{\theta} \right)^2 \quad (2.43)$$

L'estimateur de Bayes

$$\hat{\theta}_3 = \frac{E(\theta^{-1}/X)}{E(\theta^{-2}/X)} \quad (2.44)$$

Estimateur de Bayes pour α Nous posons comme dans les fonctions coût précédentes :

$u(\alpha, \beta) = \frac{1}{\alpha^2}$, donc

$$u_1 = \frac{\partial u(\alpha, \beta)}{\partial \alpha} = -2\alpha^{-3}$$

$$u_{11} = \frac{\partial^2 u(\alpha, \beta)}{\partial \alpha^2} = 6\alpha^{-4}$$

$$u_2 = \frac{\partial u(\alpha, \beta)}{\partial \beta} = 0$$

$$u_{22} = \frac{\partial^2 u(\alpha, \beta)}{\partial \beta^2} = 0$$

Alors

$$E\left(\frac{1}{\alpha^2}\right) = \frac{1}{\hat{\alpha}^2} + \frac{1}{2}(u_{11}\sigma_{11}) + p_1 u_1 \sigma_{11} + \frac{1}{2}(L_{30} u_1 \sigma_{11}^2) + \frac{1}{2}(L_{12} u_1 \sigma_{11} \sigma_{22})$$

$$E\left(\frac{1}{\alpha^2}\right) \approx \frac{1}{\hat{\alpha}^2} + \frac{1}{2} \left(6\hat{\alpha}^{-4} \frac{1}{\Psi'(\hat{\alpha})} \right) + \left(\frac{a-1}{\hat{\alpha}} - b \right) (-2\hat{\alpha}^{-3}) \frac{1}{n\Psi'(\hat{\alpha})} \quad (2.45)$$

$$+ \frac{1}{2} \left(n\Psi''(\hat{\alpha}) 2\hat{\alpha}^{-3} \frac{1}{(n\Psi'(\hat{\alpha}))^2} \right) + \frac{1}{2} \left(\frac{n}{\hat{\beta}^2} 2\hat{\alpha}^{-3} \frac{1}{n\Psi'(\hat{\alpha})} \frac{\hat{\beta}^2}{n\hat{\alpha}} \right) \quad (2.46)$$

$$\approx \frac{1}{\hat{\alpha}^2} + \frac{1}{\hat{\alpha}^4 n\Psi'(\hat{\alpha})} - \frac{2}{\hat{\alpha}^3 n\Psi'(\hat{\alpha})} \left(\frac{a-1}{\hat{\alpha}} - b \right) + \frac{n\Psi''(\hat{\alpha})}{\hat{\alpha}^3 (n\Psi'(\hat{\alpha}))^2} + \frac{1}{\hat{\alpha}^3 n\Psi'(\hat{\alpha})} \quad (2.47)$$

En remplaçant (2.38) et (2.44) dans (2.47) nous aurons :

$$\hat{\alpha}_3 \approx \frac{\frac{1}{\hat{\alpha}} + \frac{1}{\hat{\alpha}^3 n\Psi'(\hat{\alpha})} - \frac{1}{\hat{\alpha}^2 n\Psi'(\hat{\alpha})} \left(\frac{a-1}{\hat{\alpha}} - b \right) + \frac{1}{2} \left(\frac{n\Psi''(\hat{\alpha})}{\hat{\alpha}^2 (n\Psi'(\hat{\alpha}))^2} \right) + \left(\frac{1}{2\hat{\alpha}^3 n\Psi'(\hat{\alpha})} \right)}{\frac{1}{\hat{\alpha}^2} + \frac{1}{\hat{\alpha}^4} n\Psi'(\hat{\alpha}) - \frac{2}{\hat{\alpha}^3 n\Psi'(\hat{\alpha})} \left(\frac{a-1}{\hat{\alpha}} - b \right) + \frac{n\Psi''(\hat{\alpha})}{\hat{\alpha}^3 (n\Psi'(\hat{\alpha}))^2} + \frac{1}{\hat{\alpha}^3 n\Psi'(\hat{\alpha})}} \quad (2.48)$$

Estimateur de Bayes pour β Posons que $u(\alpha, \beta) = \frac{1}{\beta^2}$, donc

$$u_1 = \frac{\partial u(\alpha, \beta)}{\partial \alpha} = 0$$

$$u_{11} = \frac{\partial^2 u(\alpha, \beta)}{\partial \alpha^2} = 0$$

$$u_2 = \frac{\partial u(\alpha, \beta)}{\partial \beta} = -2\beta^{-3}$$

$$u_{22} = \frac{\partial^2 u(\alpha, \beta)}{\partial \beta^2} = 6\beta^{-4}$$

Alors

$$E\left(\frac{1}{\beta^2}\right) \approx \frac{1}{\hat{\beta}^2} + \frac{1}{2}(u_{22}\sigma_{22}) + p_2 u_2 \sigma_{22} + \frac{1}{2}(L_{03} u_2 \sigma_{22}^2) + \frac{1}{2}(L_{21} u_2 \sigma_{11} \sigma_{22})$$

$$E\left(\frac{1}{\beta^2}\right) \approx \frac{1}{\hat{\beta}^2} + \frac{1}{2} \left(\frac{6\hat{\beta}^2}{\hat{\beta}^2 n\hat{\alpha}} \right) + \frac{2c\hat{\beta}^2}{\hat{\beta}^3 n\hat{\alpha}} + \frac{1}{2} \left(\frac{2n\hat{\alpha}^2 - 2}{\hat{\beta}^3} \frac{\hat{\beta}^4}{\hat{\beta}^3 n^2 \hat{\alpha}^2} \right) \quad (2.49)$$

$$\approx \frac{1}{\hat{\beta}^2} + \frac{2c}{\hat{\beta}^2 n\hat{\alpha}} + \frac{1}{\hat{\beta}^2 n\hat{\alpha}} \quad (2.50)$$

En remplaçant (2.50) et (2.40) dans (2.44), nous obtenons :

$$\hat{\beta}_3 = \frac{\frac{1}{\hat{\beta}} + \frac{c}{n\hat{\alpha}}}{\frac{1}{\hat{\beta}^2} + \frac{2c}{\hat{\beta}n\hat{\alpha}} + \frac{1}{\hat{\beta}^2n\hat{\alpha}}} \quad (2.51)$$

Où $\hat{\beta}$ et $\hat{\alpha}$ sont les estimateurs de maximum de vraisemblance

CHAPITRE 3

APPLICATION DE LA LOI GAMMA À LA FIABILITÉ

3.1 Introduction

La fiabilité, en tant qu'attribut humain, a été louée depuis très longtemps. Cependant, le concept de fiabilité n'a été appliqué aux systèmes techniques que depuis environ 60 ans. Il a émergé avec une signification technologique après les deux guerres mondiales et a été utilisé ensuite en relation avec la comparaison de la sécurité opérationnelle des avions.

Dans les années 1970, l'intérêt pour les aspects de risque et de sécurité liés à la construction et à l'exploitation des centrales nucléaires a augmenté, aux États-Unis comme ailleurs dans le monde. Aux États-Unis, une grande commission de recherche, dirigée par le professeur Norman Rasmussen, a été mise en place pour analyser le problème. Le projet de plusieurs millions de dollars a abouti au rapport Rasmussen, WASH-1400 (NUREG-75/014). Malgré ses faiblesses, ce rapport représente la première analyse sérieuse de la sécurité d'un système aussi complexe qu'une centrale nucléaire.

Des travaux similaires ont également été menés en Europe et en Asie. Dans la plupart des industries, de nombreux efforts sont actuellement consacrés à l'analyse des problèmes de risque et de fiabilité. Il en va de même en Norvège, en particulier dans l'industrie pétrolière offshore. Le développement pétrolier et gazier offshore en mer du Nord progresse actuellement dans des eaux de plus en plus profondes et hostiles, et un nombre croissant de systèmes de production sous-marins télécommandés sont mis en service. L'importance de la fiabilité des systèmes sous-marins présente de nombreuses similitudes avec celle des engins

spatiaux. Une faible fiabilité ne peut pas être compensée par une maintenance extensive.

Une histoire plus détaillée la technologie de fiabilité est présentée par exemple, par Knight [13] et Villemeur [26].

3.2 Notions générales

- Un système est un ensemble de composants en interaction destiné à accomplir une tâche donnée. C'est le cas par exemple des systèmes de production, systèmes de transport, systèmes informatiques, etc...
- La sûreté de fonctionnement (SdF, en anglais dependability) d'un système est la propriété qui permet à ses utilisateurs de placer une confiance justifiée dans le service qu'il leur délivre. On dit aussi que la SdF est la science des défaillances.
- Un système subit une défaillance quand il ne peut plus délivrer le service attendu.
- La panne est l'état du système résultant d'une défaillance.
- La sûreté de fonctionnement comprend 4 composantes : la fiabilité, la disponibilité, la maintenabilité et la sécurité.
 1. La fiabilité (reliability) est la caractéristique du système exprimée par la probabilité qu'il délivre le service attendu dans des conditions données et pendant une durée déterminée. La fiabilité exprime l'aptitude à la continuité du service (ex : envoyer un robot sur Mars).
 2. La disponibilité (availability) est exprimée par la probabilité que le système délivre le service attendu dans des conditions données et à un instant donné. La disponibilité caractérise donc l'aptitude du système à fonctionner quand on a besoin de lui (ex : avoir du réseau quand on veut téléphoner).
 3. La maintenabilité (maintainability) caractérise l'aptitude du système à être réparé quand il est défaillant, ou à évoluer.
 4. La sécurité (safety) caractérise l'aptitude du système à ne pas encourir de défaillances catastrophiques.
- Un système non réparable est un système qui est mis au rebut dès qu'il tombe en panne. C'est le cas des petits systèmes (par exemple des ampoules) ou des systèmes qui coûtent plus cher à réparer qu'à remplacer.
- Un système réparable est un système qui, après sa défaillance, peut être remis en état de marche par des actions de réparation ou maintenance. C'est le cas de quasiment tous les systèmes complexes (véhicules, usines, etc.).

-
- La maintenance des systèmes est essentiellement de deux types :
 - La maintenance corrective (MC) ou réparation est effectuée suite à une défaillance et a pour but de remettre le système en état de fonctionner.
 - La maintenance préventive (MP) est effectuée alors que le système fonctionne et a pour but de ralentir le vieillissement pour retarder l'occurrence des défaillances futures.

3.3 Mesures de fiabilité

Les mesures de fiabilité sont différentes suivant que les systèmes concernés sont réparables ou non réparables.

3.3.1 Mesures pour systèmes non réparables

Comme nous l'avons défini plus haut, un système non réparable est un système qui est mis au rebut dès qu'il tombe en panne. Les considérations sur les réparations ou corrections n'ont donc pas lieu d'être ici. Le seul point important est la date de panne, appelée aussi instant de défaillance, durée de vie ou durée de bon fonctionnement du système.

La durée de fonctionnement d'un appareil est le temps cumulé de fonctionnement depuis la mise en service, jusqu'à la première panne. En général l'appareil n'a pas été sollicité en permanence : le fonctionnement peut avoir été interrompu par des périodes d'inactivité dont nous ne tenons pas compte. La modélisation consiste à voir les durées de fonctionnement de chaque exemplaire comme autant de réalisations d'une variable aléatoire notée X , dont on doit déterminer la loi. En fiabilité, nous caractérisons la loi d'une durée de fonctionnement de plusieurs manières différentes.

Nous supposons toujours que X admet une densité f , continue et strictement positive sur \mathbb{R}^{+*} . La fonction de répartition F est la primitive de la densité, nulle en 0. En plus de la densité et de la fonction de répartition, habituellement utilisées pour caractériser la loi d'une variable continue, nous introduisons deux nouvelles fonctions liées aux précédentes, la fiabilité et le taux de défaillance (ou de panne), ainsi que les temps moyens.

Définition 16. *Nous appelons fiabilité d'une durée de fonctionnement X , le complément à 1 de la fonction de répartition de X . Cette fonction, notée R (pour "Reliability"),*

vérifie pour tout $t \geq 0$:

$$\forall x \geq 0 \quad R(t) = \mathbb{P}[X > t] = 1 - F(t) = \int_t^{+\infty} f(s)ds \quad (3.1)$$

La fiabilité est la probabilité que l'appareil fonctionne au-delà de l'instant t . Nous l'appelons aussi fonction de survie. C'est une fonction continue, strictement positive et décroissante, qui vaut 1 en $t = 0$ et qui tend vers 0 quand t tend vers l'infini.

Définition 17. *Le taux de défaillance ou taux de panne ou taux de hasard d'un système non réparable est la fonction du temps h définie par :*

$$\forall x \geq 0, \quad h(x) = \lim_{\Delta x \rightarrow 0} \frac{1}{\Delta x} \mathbb{P}(x < X \leq x + \Delta x | X > x) \quad (3.2)$$

Définition 18. *Le taux de défaillance cumulé ou taux de hasard cumulé d'un système non réparable est la fonction du temps H définie par :*

$$\forall x \geq 0, \quad H(x) = \int_0^x h(u)du = -\ln R(x) \quad (3.3)$$

Définition 19. *Nous appelons MTTF (Mean Time To Failure) : la durée moyenne de bon fonctionnement du composant/appareil. Il est défini comme suit :*

$$MTTF = E(T) = \int_0^{+\infty} \mathbb{P}[X > t]dt = \int_0^{+\infty} R(t)dt \quad (3.4)$$

3.3.2 Mesures pour systèmes réparables

Quand les systèmes sont réparables, deux cas de figure sont possibles, selon que l'on prend en compte ou pas les durées de réparation.

Durées de réparation comptabilisées

Le fonctionnement du système est une succession de durées de bon fonctionnement et de durées de non fonctionnement ou de réparation. Nous notons généralement $\{X_i\}_{i \geq 1}$ durées de bon fonctionnement successives et $\{Y_i\}_{i \geq 1}$ les durées de réparation successive.

La "durée de réparation" Y comprend en fait une durée de détection de la panne, une durée de réparation proprement dite et une durée de remise en service. Pour une durée de réparation Y , nous définissons les quantités similaires à celles qui ont été définies pour une durée de bon fonctionnement X :

Définition 20. La maintenabilité est la fonction de répartition de Y . La maintenabilité en y est la probabilité qu'un système en panne à l'instant 0 soit réparé avant l'instant y :

$$\forall y \geq 0, \quad M(y) = \mathbb{P}[Y \leq y] \quad (3.5)$$

Définition 21. Le taux de réparation est défini par :

$$\forall y \geq 0, \quad \mu(x) = \lim_{\Delta y \rightarrow 0} = \frac{1}{\Delta y} \mathbb{P}(y < Y \leq y + \Delta y | Y > y) \quad (3.6)$$

Définition 22. Le MTTR (Mean Time To Repair) est la durée moyenne de réparation :

$$MTTR = \mathbb{E}[Y] = \int_0^{+\infty} [1 - M(y)] dy \quad (3.7)$$

Durées de réparation non comptabilisées

En pratique, il est fréquent que les durées de réparation soient négligeables par rapport aux durées de bon fonctionnement. Il est donc intéressant de modéliser la situation où les durées de réparation sont non comptabilisées. Dans ce cas, la notion de disponibilité n'a plus aucun sens.

Dans ces conditions, nous considérons que nous observons le fonctionnement d'un système réparable à partir d'un instant $T_0 = 0$. Des défaillances se produisent à des instants que l'on note T_1, T_2, \dots . Après chaque défaillance, le système est réparé ou corrigé puis relancé. Le processus des défaillances d'un tel système réparable est défini de manière équivalente par l'un des 3 processus aléatoires suivants :

- La suite des instants de défaillance $\{T_i\}_{i \geq 1}$, avec $T_0 = 0$
- La suite des durées inter-défaillances $\{X_i\}_{i \geq 1}$ où $\forall i \geq 1, X_i = T_i - T_{i-1}$ est la durée entre la $(i-1)$ ème et i ème défaillance. $T_i = \sum_{j=1}^i X_j$
- Le processus de comptage des défaillances $\{N_t\}_{t \geq 0}$ où N_t est le nombre cumulé de défaillances survenues entre 0 et t .

Définition 23. La fiabilité d'un système réparable à l'instant t , ayant subi n défaillances avant t , est la fonction R_t définie par :

$$\forall \tau \geq 0, \quad R_t(\tau; n, t_1, \dots, t_n) = \mathbb{P}(T_{n+1} > t + \tau | N_t = n, T_1 = t_1, \dots, T_n = t_n) \quad (3.8)$$

$$= \mathbb{P}(N_t + \tau - N_t = 0 | N_t = n, T_1 = t_1, \dots, T_n = t_n) \quad (3.9)$$

Définition 24. *Le MTTF d'un système réparable à l'instant t est la durée moyenne d'attente de la prochaine défaillance à l'instant t , sachant tout le passé du processus des défaillances à cet instant :*

$$MTTF_t(n; t_1, \dots, t_n) = \mathbb{E}[T_{n+1} - t | N_t = n, T_1 = t_1, \dots, T_n = t_n] \quad (3.10)$$

Définition 25. *La fonction moyenne (en anglais mean value function) du processus des défaillances est la fonction m définie par :*

$$\forall t \geq 0, m(t) = \mathbb{E}[N_t] \quad (3.11)$$

3.3.3 Lois classiques utilisées dans la fiabilité

Il y'a trois familles de lois essentielles pour la fiabilité : la famille Exponentielle, la famille Weibull et la famille Gamma. Nous allons nous focaliser sur la loi Gamma, car c'est celle traitée dans ce mémoire. La loi Gamma a été définie au premier chapitre.

La fiabilité et le taux de défaillance de la loi Gamma $\Gamma(\alpha, \lambda)$ s'expriment à l'aide de la fonction gamma incomplète, définie par :

$$\Gamma(\alpha, t) = \int_t^{+\infty} e^{-s} s^{\alpha-1} ds$$

Cette fonction est calculable numériquement, et codée dans la plupart des environnements mathématiques. On démontre que le taux de défaillance de la loi $\Gamma(a, \lambda)$ converge vers λ quand t tend vers l'infini. Pour $a < 1$ c'est une fonction décroissante de t , qui tend vers $+\infty$ en $t = 0$. Pour $a > 1$, c'est une fonction croissante de t qui tend vers 0 en $t = 0$.

- Si $\alpha < 1$, h est décroissant donc le système s'améliore ;
- Si $\alpha > 1$, h est croissant donc le système s'use ;
- Si $\alpha = 1$, h est constant et nous retrouvons la loi exponentielle. Nous retrouvons ces trois cas représentés dans la figure suivante :

3.4 Censure

Nous parlons de données censurées quand il peut être trop cher, inutile, ou tout simplement impossible, d'observer toutes les durées de fonctionnement x_1, \dots, x_n . Nous décidons alors d'interrompre l'observation avant la dernière. Nous ne considérerons que deux types de censure. Il existe de nombreuses généralisations de ces deux cas de base :

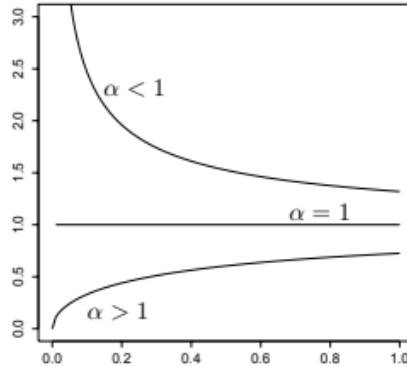


FIGURE 3.1 – Taux de défaillance d’une loi Gamma

3.4.1 Type I

Les données n’ont été observées que jusqu’à un instant de censure t_c fixé à l’avance. Dans ce cas, l’observation se compose des valeurs de celles des durées qui étaient inférieures à t_c . En d’autres termes, au lieu d’observer la valeur prise par X_i , nous observons celle de X'_i , avec :

$$X'_i = X_i \mathbf{1}_{[0, t_c]}(X_i) + t_c \mathbf{1}_{]t_c, +\infty[}(X_i)$$

La variable X'_i vaut t_c avec probabilité $R(t_c)$, elle est égale à X_i si $X_i \geq t_c$. La probabilité de l’événement observé est proportionnelle à :

$$\mathbf{L} = \prod_{i=1}^n (f(x_i) \mathbf{1}_{[0, t_c]}(x_i) + R(t_c) \mathbf{1}_{]t_c, +\infty[}(x_i)) \quad (3.12)$$

Pour écrire cette formule sous une forme plus lisible, notons $x_{(1)}, \dots, x_{(n)}$ les valeurs x_1, \dots, x_n rangées par ordre croissant, et k le nombre de valeurs inférieures à t_c :

$$k = \sum_{i=1}^n \mathbf{1}_{[0, t_c]}(x_i) = \max(i, x_{(i)} \leq t_c)$$

Les variables aléatoires $X_{(1)}, \dots, X_{(n)}$ (valeurs de X_1, \dots, X_n rangées par ordre croissant) s’appellent les statistiques d’ordre de l’échantillon.

La vraisemblance devient :

$$\mathbf{L} = R(t_c)^{n-k} \prod_{i=1}^k (f(x_{(i)})) \quad (3.13)$$

3.4.2 Type II

Seules les n_c premières valeurs ont été observées, où n_c est un nombre fixé à l'avance. Comme précédemment, notons $x_{(1)}, \dots, x_{(n)}$ les valeurs rangées par ordre croissant. Nous observons en fait les valeurs prises par $X_{(1)}, \dots, X_{(n_c)}$, les n_c premières statistiques d'ordre de l'échantillon. La probabilité de l'événement observé est ici proportionnelle à :

$$\mathbf{L} = R(x_{(n_c)})^{n-n_c} \prod_{i=1}^{n_c} (f(x_{(i)})) \quad (3.14)$$

3.5 Application de la loi Gamma à la fiabilité

Beaucoup de grandes organisations tels que l'US Air Force et la Navy créent des bases de données pour traquer la fiabilité des systèmes et composants sur le terrain. Le "Reliability Analysis Center (RAC)" (Le center d'analyse de fiabilité en français) est un centre de recherche du département de la défense de l'état américain dédié à la récolte et la sauvegarde des données liées à la fiabilité, c'est la plus grande source public de données liées à la fiabilité disponible. Cependant, pour une variété de raisons 90% de leurs données n'ont pas de temps de défaillance individuels, mais les temps de défaillance cumulés sont disponibles. Alors, comme les temps de défaillance ne sont pas fidèlement enregistrés (il y'a des temps de défaillances manquants) ceci résulte en une forme non standard de données aléatoires qui ne peut être étudiée par les techniques d'étude de fiabilité disponibles.

Avantages et inconvénients des données réelles

Pour évaluer la fiabilité d'un composant, les données sur le terrain présentent de nombreux avantages distincts. L'avantage principal est que le stress et les contraintes sont appliquées simultanément, et les interactions variables sont implicitement prises en compte.

Mais, il y'a aussi quelques inconvénients aux données réels, en effet dans quelques bases de données les temps de défaillance individuels sont souvent manquants. Les données ne sont souvent disponibles que sous forme de défaillances collectives observées en cumulant des heures sans autre détail disponible.

Notations et hypothèses

Les hypothèses suivantes ont été faites concernant la défaillance des composants et la collection des données du système.

- Les temps de défaillance des composants sont iid.
- Les temps de défaillance sont distribués en accord avec une loi Gamma.
- Les temps de réparation sont insignifiants par rapport aux temps de fonctionnement.
- La réparation du système ne se dégrade pas et n'affecte pas la fiabilité des composants.

Les notations suivantes sont utilisées dans la construction de l'estimateur de maximum de vraisemblance de la loi Gamma.

- λ, k = les paramètres de la loi gamma $f(t) = \lambda^k t^{k-1} e^{-\lambda t} / \Gamma(k)$;
- $\hat{\lambda}, \hat{k}$ = estimateurs du maximum de vraisemblance de λ, k ,
- r = le nombre de défaillances observées;
- T_{rj} = j^{ème} cumulation des temps d'opération avec r défaillances = $X_1 + X_2 + \dots + X_r$;
- X_i = i^{ème} temps de défaillance;
- n_r = nombre de données enregistrées avec exactement r défaillances;
- $n = (n_1, n_2, \dots, n_m)$;
- m = le nombre max de défaillances pour n'importe quelle donnée enregistrée dans un lot de données = $\max(r | n_r > 0)$;
- N = nombre de données enregistrées, $N = \sum_{r=1}^m n_r$;
- M = nombre total de défaillances associées avec N toutes les données enregistrées, $M = \sum_{r=1}^m r n_r$;
- \bar{t} = temps moyen avant la panne = $\sum_{r=1}^m \sum_{j=1}^{n_r} T_{rj} / M$;
- $\Gamma(z)$ = fonction Gamma = $\int_0^\infty u^{z-1} e^{-u} du$.

3.5.1 Application sur des données réelles

Les auteurs "DAVID W. COTT" et "TONGDAN JIN" dans l'article "Gamma distribution parameter estimation for field reliability data with missing failure times" [5], en 1999, se sont intéressés aux bases de données du RAC, et ont pris un exemple de fiabilité un système de composant qui sont des voyants lumineux d'avion.

En effet comme nous remarquons que les temps individuels de défaillance ne sont pas disponibles, mais les temps cumulés eux le sont.

Le tableau 3.1 présente l'ensemble de données que les auteurs ont choisis.

Défaillances	Temps d'opération cumulés(heures)
2	$T_{21} = 51000$
9	$T_{91} = 194900$
8	$T_{81} = 45300$
8	$T_{82} = 112400$
6	$T_{61} = 104000$
5	$T_{51} = 44800$

TABLE 3.1 – Données de fiabilité de voyants lumineux d'avion

Temps de défaillances manquants

En général, les compteurs enregistrent les heures cumulées de fonctionnement (ou une autre mesure) d'un système assemblé par opposition à un composant individuel au sein du système.

Les temps de défaillance au niveau des composants sont ensuite déterminées comme la différence entre le temps du fonctionnement du système lorsque le composant est installé et la durée de fonctionnement du système à la panne.

Les temps de fonctionnement ne sont pas toujours enregistrés par les opérateurs et les techniciens de maintenance, donc une estimation précise des paramètres de la distribution des temps de défaillance est nécessaire pour déterminer les mesures préventives, la période de garantie ..

Lorsque le temps de fonctionnement d'un système associé à une défaillance sur le terrain d'un composant est observé et enregistré avec précision, la seule information qui peut être connue est le nombre total de pannes de ce composant depuis le précédent temps de défaillance enregistré. Les temps intermédiaires avant le temps de défaillance ne sont pas connus bien que le nombre d'événements de défaillance l'est.

Comme les temps de défaillances des composants sont manquants (inconnus), il n'y a pas de solution analytique.

Les lois des temps de défaillance

Pour les ensembles de données standard, il existe des techniques bien connues pour estimer les paramètres de nombreuses distributions et types de censure, et pour évaluer objectivement l'applicabilité de ces distributions. Cependant, sans données sur les temps de défaillance individuels, il est impossible d'ajuster les données à la plupart des distributions populaires (par exemple, Weibull, exponentielle, ect ...) en utilisant des techniques standard telles que l'estimation ou l'analyse de

régression. C'est pour cela que l'auteur, a choisit l'estimateur du maximum de vraisemblance de la loi Gamma.

L'estimateur de maximum de vraisemblance

La distribution des temps de défaillance cumulées L'approche pour déterminer l'estimateur de maximum de vraisemblance de la distribution Gamma de paramètres (λ, k) est de développer la fonction de vraisemblance basée sur les paramètres de la loi (λ, k) , les r observés et les valeurs de T_{rj} au lieu des temps de défaillance et les censures (qui sont inconnues). Quand X_i suit une loi Gamma, alors T_{rj} suit lui aussi une loi gamma pour un r fixé.

- $X_i =$ temps de défaillance du composant;
- $X_i \sim \Gamma(\lambda, k), E(X) = k/\lambda, Var(X) = k/\lambda^2$;
- $T_{rj} = \sum_{i=1}^r X_i$;
- $T_{rj} \sim \Gamma(rk, \lambda), E(X) = k/\lambda, Var(X) = k/r\lambda^2$.

T_{rj} est une variable aléatoire dépendant d'un r spécifique. Si r est aussi une variable aléatoire, considérons alors T comme le temps de défaillance cumulé associé pour un r aléatoire. La fonction de densité de T_{rj} est alors une densité de probabilité conditionnelle de T . Par ailleurs, la fonction de densité de T peut-être exprimée par une somme de termes conditionnels.

$$f_{T_r}(t) = f_T(t|r) = \frac{\lambda^{rk} t^{rk-1} \exp(-\lambda t)}{\Gamma(rk)} \quad (3.15)$$

$$f_T(t) = \sum_{r=1}^{\infty} f_T(t|r) Pr(R = r)$$

- Où $Pr(R = r)$ est la probabilité d'avoir r échecs.
- Considérons un ensemble de N données enregistrées avec r et T_{rj} connus. L'ensemble des temps cumulés de fonctionnement (T_{rj}) forme une population non homogène qui peut être modélisée à l'aide d'un modèle de mixture.
- Les variables T_{rj} sont associées à différentes distributions (dépendant de r) mais elles sont liées à la Gamma.
- Les modèles de mixture sont généralement utilisées lorsque la population n'est pas homogène(elle est composée de sous-population distinctes).
- La situation la plus courante est celle où la proportion de chaque population est connue ou estimée.

-
- Un exemple courant est celui de la fabrication de semi-conducteur où il existe une sous-population d'appareils faibles ou défectueux mais nous ne pouvons pas distinguer entre les deux sous-populations avant d'introduire les composants sur le terrain.
 - La fonction de densité pour une population non homogène peut être exprimée comme une somme pondérée des densités respectives.
 - Pour le problème abordé ici, les sous-populations sont caractérisées par le nombre d'échecs au sein de l'ensemble de données
 - La probabilité que $T_{r,j}$ sélectionnée au hasard parmi les n enregistrés de données à r échecs est de probabilité n_r/N est :

$$f_T(t|n) = \frac{1}{N} \sum_{r=1}^m \frac{n_r \lambda^{rk} t^{rk-1} e^{-\lambda t}}{\Gamma(rk)}. \quad (3.16)$$

Passons maintenant à l'estimateur. Une fonction de vraisemblance pour λ et k peut-être présentée directement basé sur les données observées et l'équation (3.16).

$$\mathbf{L} = \prod_{r=1}^m \left\{ \prod_{j=1}^{n_r} \left[\frac{1}{N} \frac{n_i \lambda^{ik} T_{rj}^{ik-1} e^{-\lambda T_{rj}}}{\Gamma(ik)} \right] \right\} \quad (3.17)$$

Nous obtenons les estimateurs de λ et k en utilisant la méthode de Newton-Raphson (en minimisant \mathbb{L}). Voici les étapes suivis :

$$\mathbf{L} = \prod_{r=1}^m \prod_{j=1}^{n_r} f_T(t|n) = \prod_{r=1}^m \prod_{j=1}^{n_r} \left[\frac{1}{N} \sum_{i=1}^m \frac{n_i \lambda^{ik} T_{rj}^{ik-1} e^{-\lambda T_{rj}}}{\Gamma(ik)} \right]$$

Une autre formule de la fonction de vraisemblance peut-être développée en exploitant les observations de la sous-population, cette fonction de vraisemblance alternative peut-être exprimée comme le produit des fonctions de densités conditionnelles pour $T_{r,j}$ ($f_T(t|r)$).

$$f_T(t|r) = \frac{\lambda^{rk} t^{rk-1} e^{-\lambda t}}{\Gamma(rk)}$$

$$\begin{aligned}
L(k, \lambda) &= \prod_{r=1}^m \prod_{j=1}^{n_r} \frac{\lambda^{rk} T_{rj}^{rk-1} e^{-\lambda T_{rj}}}{\Gamma(rk)} \\
&= \prod_{r=1}^m \prod_{j=1}^{n_r} \frac{\lambda^{rk} T_{rj}^{rk-1}}{\Gamma(rk)} \left(\exp \left(-\lambda \sum_{r=1}^m \sum_{j=1}^{n_r} T_{rj} \right) \right) \\
&= \prod_{r=1}^m \left(\prod_{j=1}^{n_r} \left[\frac{1}{N} \sum_{i=1}^m \frac{n_i \lambda^{ik} T_{rj}^{ik-1} e^{-\lambda T_{rj}}}{\Gamma(ik)} \right] \right)
\end{aligned}$$

Avec $X_i \sim \Gamma(\lambda, k)$, nous avons $T_{rj} = \sum_{i=1}^r X_i = x_1 + x_2 + \dots + x_r$ donc $T_{rj} \sim \Gamma(\lambda, rk)$

$$\begin{aligned}
f_{T_r}(t) &= f_r(t|r) = \frac{\lambda^{rk} t^{rk-1} e^{-\lambda t}}{\Gamma(rk)} \sim \Gamma(\lambda, rk) \\
f_T(t) &= \sum_{r=1}^{\infty} f_r(t|r) \cdot Pr(R = r) \\
f_T(t) &= \sum_{r=1}^{\infty} \frac{\lambda^{rk} t^{rk-1} e^{-\lambda t}}{\Gamma(rk)} \cdot Pr(R = r)
\end{aligned}$$

Nous savons que $Pr(R = r) = \frac{n_r}{N}$ alors :

$$f_T(t) = \frac{1}{N} \sum_{r=1}^{\infty} \frac{n_r \lambda^{rk} t^{rk-1} e^{-\lambda t}}{\Gamma(rk)}$$

La fonction devient alors :

$$f_T(t|n) = \frac{1}{N} \sum_{r=1}^{\infty} \frac{n_r \lambda^{rk} t^{rk-1} e^{-\lambda t}}{\Gamma(rk)} \quad (3.18)$$

Passons maintenant à la log-vraisemblance :

$$\begin{aligned}
\ln(\mathbf{L}(k, \lambda)) &= -\lambda \sum_{r=1}^m \sum_{j=1}^{n_r} T_{rj} + \sum_{r=1}^m \sum_{j=1}^{n_r} rk \ln(\lambda) + k \sum_{r=1}^m \sum_{j=1}^{n_r} k \ln(T_{rj}) \\
&\quad - \sum_{r=1}^m \sum_{j=1}^{n_r} \ln(T_{rj}) - \sum_{r=1}^m \sum_{j=1}^{n_r} \ln(\Gamma(rk))
\end{aligned}$$

$$\ln(\mathbf{L}(k, \lambda)) = -\lambda \bar{t} M + M k \ln \lambda + k \sum_{r=1}^m \sum_{j=1}^{n_r} r \ln T_{rj} - \sum_{r=1}^m \sum_{j=1}^{n_r} \ln T_{rj} - \sum_{r=1}^m n_r \ln(\Gamma(rk)) \quad (3.19)$$

Les dérivées partielles par rapport à λ et k sont :

Nous savons que $M = \sum_{r=1}^m r n_r$, alors :

$$\frac{\partial \ln(\mathbf{L}(k, \lambda))}{\partial \lambda} = -M \bar{t} + \frac{Mk}{\lambda} \quad (3.20)$$

$$\frac{\partial \ln(\mathbf{L}(k, \lambda))}{\partial k} = M \ln \lambda + \sum_{r=1}^m \sum_{j=1}^{n_r} r \ln T_{rj} - \sum_{r=1}^m r n_r \Psi(rk) \quad (3.21)$$

Posons aussi :

$$\Psi(rk) = \frac{\Gamma'(rk)}{\Gamma(rk)} \quad (3.22)$$

De ((3.20)) nous avons :

$$\hat{\lambda} = \frac{\hat{k}}{\bar{t}} \quad (3.23)$$

En mettant ((3.21)) = 0 nous aurons

$$\begin{aligned} \sum_{r=1}^m \sum_{j=1}^{n_r} r \ln T_{rj} &= \sum_{r=1}^m r n_r \Psi(rk) - M \ln(k|\bar{t}) \\ \sum_{r=1}^m \sum_{j=1}^{n_r} r \ln T_{rj} - M \ln(\bar{t}) &= \sum_{r=1}^m r n_r \Psi(rk) - M \ln(k) \end{aligned}$$

Ce qui nous donne

$$\hat{k} = \left\{ k \left| \sum_{r=1}^m r n_r \Psi(rk) - M \log k = K' \right. \right\} \quad (3.24)$$

Avec $K' = \sum_{r=1}^m \sum_{j=1}^{n_r} r \ln T_{rj} - M \ln \bar{t}$

Avec \hat{k} et $\hat{\lambda}$ les estimateurs de maximum de vraisemblance.

Simulation

Grâce au premier exemple étant celui des données réels du tableau 3.1 et aux formules 3.24 et 3.23, où $(m = 9, N = 6, M = 38, n = (0, 1, 0, 0, 1, 1, 0, 2, 1))$, nous avons pu calculer \bar{t} , k et λ , cependant les autres exemples ont été simulés avec le logiciel R.

Nous savons aussi que \hat{k} est la solution de l'équation 3.25 $f(k) = 0$ qui est une équation non linéaire.

$$f(k) = \sum_{r=1}^m r n_r \Psi(rk) - M \ln k - \sum_{r=1}^m \sum_{j=1}^{n_r} r \ln T_{rj} + M \ln \bar{t} = 0 \quad (3.25)$$

Exemple 2 ($k = 1, \lambda = 0.00006879$)		Exemple 3 ($k = 3, \lambda = 0.0002064$)	
Défaillances	T_{rj}	Défaillances	T_{rj}
2	28 131	2	46 170
5	61 363	5	83 170
6	64 995	6	88 950
8	98 859	8	110 530
8	145 683	8	103 210
9	37 607	9	93 010

TABLE 3.2 – Données pour les exemples simulés

Exemple 4 ($k = 2, \lambda = 0.00013758$)		Exemple 5 ($k = 0.8, \lambda = 0.0005503$)	
Défaillances	T_{rj}	Défaillances	T_{rj}
2	40 173	2	74 112
5	118 749	5	267 520
6	98 817	6	213 790
8	109 306	8	345 125
8	43 833	8	92 669
9	72 586	9	68 028

TABLE 3.3 – Données pour les exemples simulés

Les estimateurs de maximum de vraisemblance trouvés précédemment sont dans le tableau 3.4 pour chaque exemple

Le numéro de l'exemple	\hat{k}	$\hat{\lambda}$
1	0.70	0.000 0484
2	0.70	0.000 0609
3	3.43	0.000 2483
4	1.82	0.000 1478
5	1.09	0.000 6726

TABLE 3.4 – Résultats des exemples

3.6 Conclusion

Les bases de données disponibles sont souvent incomplètes et ceci provient du système de collecte de données non rigoureux où les délais de défaillances des composants ne sont pas toujours enregistrés pour distinctes raisons. En général les praticiens utilisent les méthodes standards comme le modèle exponentielle et de Weibull, cependant elles ne peuvent pas marcher dans un cas comme celui traité dans l'article.

C'est pour cela que l'auteur a proposé une approche plus rigoureuse, celle d'un estimateur de maximum de vraisemblance d'une distribution Gamma sur la base d'enregistrements des temps de défaillance qui manquent, car la loi gamma peut modéliser divers comportements en matière de temps de défaillance, elle fournit un outil pour les données qui ne pourraient pas être analysées de manière satisfaisante.

Mais, nous remarquons que les estimateurs de maximum de vraisemblance sont malheureusement biaisés, et pour cela il serait intéressant d'améliorer ce travail en utilisant l'estimation Bayésienne de la loi Gamma.

CONCLUSION GÉNÉRALE

Dans ce mémoire nous avons examiné en détail la distribution gamma, ses propriétés fondamentales, ses applications diverses et ses extensions. Notre exploration approfondie de cette distribution a révélé son importance dans un large éventail de domaines. Nous avons mis en lumière ses aspects théoriques, et ses applications pratiques à travers un exemple concret.

En envisageant les perspectives futures, il est clair que la distribution gamma continuera de jouer un rôle majeur dans la modélisation statistique et la résolution de problèmes pratiques. En prenant exemple de l'application que nous avons faite dans ce mémoire, nous pourrions utiliser une estimation bayésienne sous différentes fonctions coûts à fin d'estimer les temps de défaillance manquants.

Enfin, les avancées technologiques en matière de collecte et d'analyse de données ouvriront de nouvelles opportunités pour son utilisation. De plus, des recherches futures pourraient se concentrer sur la compréhension des interactions entre la loi gamma et d'autres distributions statistiques, ainsi que sur le développement de méthodes d'estimation plus avancées.

.1 La Méthode de Lindley

Dennis Lindley a développé la procédure d'approximation, pour les intégrales de la forme suivante en 1980.

$$\int \frac{w(\theta) \exp\{l(\theta)\} d\theta}{v(\theta) \exp\{l(\theta)\} d\theta} \quad (26)$$

Avec $\theta = (\theta_1, \theta_2, \dots, \theta_m)$, $l(\theta) = \log L(\theta|x)$ est le logarithme de la fonction de vraisemblance.

$w(\theta)$ et $v(\theta)$ sont des fonctions arbitraires de θ .

Si $v(\theta)$ est la densité à priori de θ et $w(\theta) = \Phi(\theta) \cdot v(\theta)$, alors l'équation (26) donne l'espérance a posteriori de $\Phi(\theta)$, telle que :

$$\begin{aligned} E(\Phi(\theta)|x) &= \frac{\int \Phi(\theta) \exp\{l(\theta) + p(\theta)\} d\theta}{\int \exp\{l(\theta) + p(\theta)\} d\theta} \\ &= \frac{\int \Phi(\theta) \exp\{\Lambda(\theta)\} d\theta}{\int \exp\{\Lambda(\theta)\} d\theta} \end{aligned}$$

Avec $p(\theta) = \log v(\theta)$ et $\Lambda(\theta) = \log\{\pi(\theta|x)\} = l(\theta) + p(\theta)$ le logarithme de la distribution a posteriori de θ .

Il est évident que le maximum de $\Lambda(\theta) = \log \pi(\theta|x)$ nous donne le mode à posteriori de θ .

Lindley a donc obtenu l'expression requise pour $E(\Phi(\theta)|x)$.

Prenons maintenant $S_t = \Phi(\theta)$, en utilisant la méthode de Lindley, l'estimateur Bayésien pour S_t est :

$$S_t^* = \Phi(\hat{\theta}) + \frac{1}{2}\Sigma\Phi_{ij}(\hat{\theta})\tau_{ij} + \frac{1}{2}\Sigma\Lambda_{ijk}(\hat{\theta})\tau_{ij}\tau_{kl} \quad (27)$$

Avec

$$\Phi_{ij} = \frac{\partial^2\Phi}{\partial\theta_i\partial\theta_j}$$
$$\Lambda_{ijk} = \frac{\partial^3\Phi}{\partial\theta_i\partial\theta_j\partial\theta_k}, \dots$$

Les τ_{ij} sont les (i, j) -èmes éléments de l'inverse de la matrice Hessienne au signe négatif.

La matrice des secondes dérivées pour $\Lambda : \{\tau_{ij}\} = \{-\Lambda_{ij}\}^{-1}$, où $\hat{\theta}$ est le mode a posteriori.

BIBLIOGRAPHIE

- [1] Agarwall, a generalized gamma distribution and its application in reliability, 1996.
- [2] Bendersky, sur la fonction gamma généralisée, 1933.
- [3] Bownan K. O. and L. R. Shenton, Properties of estimators for the gamma distribution, 2007.
- [4] Choi S. C. and Wette R., Maximum Likelihood Estimation of the Parameters of the Gamma Distribution and Their Bias, 1969.
- [5] Coit W. David and Jin Tongdan, Gamma distribution parameter estimation for field reliability data with missing failure times, 1999.
- [6] École Polytechnique Universitaire Pierre et Marie Curie, Cours d'introduction, Électronique et informatique Sciences de la Terre, Probabilités, 2014.
- [7] Ferguson G.A. et Takane Y., Statistical Analysis in Psychology and education, McGraw-Hill Book, (1989)
- [8] Garson B. David, Validity and Reliability : 2016 Edition (Statistical Associates Blue Book Series 12).
- [9] Gaudoin Olivier, Fiabilité des Systèmes, Notes de cours.
- [10] Gupta Isha and Gupta Rahul, Bayesian and Non-Bayesian Method of Estimation of Scale Parameter of Gamma Distribution under Symmetric and Asymmetric Loss Functions (2018).
- [11] Khodabina, Morteza and Ahmadabadib, Alireza, Some properties of generalized gamma distribution, 2010.
- [12] Kirimi Emily, Abel Ouko, Cheruiyot W. Kipkoech, Modified Moment Estimation for a Two-Parameter Gamma Distribution, 2014.
- [13] Knight A. Gary, Cross-cultural reliability and validity of a scale to measure firm entrepreneurial orientation, 1991.

-
- [14] Lukacs Eugene, A Characterization of the Gamma Distribution, 1965.
- [15] Mihram G. Arthur and Stacy E. Webb, Parameter estimation for a generalized gamma distribution, 1965.
- [16] Miller B. Robert, Bayesian Analysis of the Two-Parameter Gamma Distribution 2012.
- [17] MOALA ANTONIO, FERNANDO and LUIZ RAMOS, PEDRO and ALBERTO ACHCAR, JORGE Bayesian inference for two-parameter gamma distribution assuming different noninformative priors, 2013.
- [18] Minka Thomas P., Estimating a Gamma distribution 2002.
- [19] Pradhan Biswabrata and Debasis Kundu, Bayes estimation and prediction of the two-parameter gamma distribution 2010.
- [20] Pradhan, Biswabrata and Kundu, Debasis, Bayes estimation and prediction of the two-parameter gamma distribution, 2011.
- [21] Rasheed A. Huda and Loaiy F. Naji, Bayesian and Non-Bayesian Parameters Estimation of Gamma distribution at Various Loss Functions.
- [22] Rasheed A. Huda and Loaiy F. Naji, Bayesian Estimation for Two Parameters of Gamma Distribution under Generalized Weighted Loss Function.
- [23] Rasheed A. Huda and Loaiy F. Naji, Bayesian estimation for two parameters of Gamma distribution under precautionary loss function 2019.
- [24] Stacy E. W., A Generalization of the Gamma Distribution, 1955.
- [25] THOM H. C. S., A note on the gamma distribution, 1958.
- [26] Villemeur, Alain ; Caseau, Paul ; D'Harcourt, Arnould, Sûreté de fonctionnement des systèmes industriels : fiabilité, facteurs humains, informatisation 1988.
- [27] Wang Bing Xing and Wu Fangtao, Inference on the gamma distribution, Technometrics, 2017.
- [28] Ycart, Cahier de Mathématiques Appliquées no 13, Notions de fiabilité.
- [29] Young Sook Son and Oh Mira, Bayesian Estimation of the Two Parameter Gamma Distribution 2007.
- [30] Zhu Xiaojun and Balakrishna N., Some simple estimators for the two-parameter gamma distribution, Helton Saulo, Marcelo Bourguignon, 2018.