

# وزارة التعليم العالي والبحث العلمي

BADJI MOKHTAR-ANNABA  
UNIVERSITY  
UNIVERSITE BADJI MOKHTAR  
ANNABA



جامعة باجي مختار  
-عناية-

Faculté des Sciences

Année : 2022/2023

Département de Mathématiques



## THÈSE

Présentée en vue de l'obtention du diplôme de Doctorat

### ANALYSE STATISTIQUE DES MODÈLES GÉNÉRALISÉS DE DURÉES DE VIE

Filière

Mathématiques Appliquées

Spécialité

Probabilités et Statistique

Par

MERIBOUT Khaoula Kaouter

DIRECTEUR DE THÈSE : SEDDIK AMEUR Nacira Prof. U.B.M. ANNABA

CO-DIRECTEUR DE THÈSE : GOUAL Hafida M.C.A. U.B.M. ANNABA

Devant le jury

PRESIDENT: CHADLI Assia Prof. U.B.M. ANNABA

EXAMINATEUR : LEULMI Sarra M.C.A. U. CONSTANTINE 1

EXAMINATEUR : MEZHOUD Kenza Assia M.C.A. U. CONSTANTINE 1

EXAMINATEUR: TALHI Hamida M.C.A. U.B.M. ANNABA

# Remerciements

Je remercie **ALLAH** mon seigneur pour toutes les grâces dont il me comble sans cesse et éclaire mon chemin, pour ses bienfaits et pour m'avoir aidé et guidé dans l'accomplissement de cette thèse. Le tout puissant proclama « **Si vous êtes reconnaissant, j'augmenterai mes bienfaits pour vous** », IBRAHIM [7], et il dit « **Souvenez- vous de Moi donc, Je Me souviendrai de vous. Remerciez-Moi et ne soyez pas ingrats envers Moi** » El- BAQARA [152], et il dit « **Lis, Ton Seigneur est le Très Noble, qui a enseigné par la plume, a enseigné à l'homme ce qu'il ne savait pas** », AL- ALAQ [3,4,5].

Je tiens à exprimer ma gratitude et à remercier en premier lieu ma directrice de thèse, Madame le Professeur **SEDDIK-AMEUR Nacira** qui n'a épargné aucun effort pour me transmettre son savoir, son expérience, son aide, ses judicieux conseils à la hauteur de ses compétences et de ses réelles qualités humaines, ainsi que sa patience, sa tolérance et son temps précieux qui ont contribué à enrichir ma réflexion et la rédaction de cette thèse.

Je tiens à témoigner toute ma reconnaissance à ma co-directrice de thèse, Docteur **GOUAL Hafida** qui m'a beaucoup appris sur le défi à relever dans le vaste monde de la Statistique, pour avoir partagé ses connaissances et son expérience en m'accordant sa confiance, son temps et une large indépendance dans l'exécution de ce travail si difficile et si valorisant.

Mes remerciements s'adressent au Président du Laboratoire des Probabilités et Statistique, Le Professeur **ZEGHDOUDI Halim** pour sa disponibilité et son aide.

Je remercie également les intervenants professionnels responsables de la formation doctorale.

J'exprime ma gratitude au Professeur **CHADLI Assia**, pour nos échanges au cours de cette thèse mais aussi pour l'hommage qu'elle me fait en présidant le jury.

Mes chaleureux remerciements s'adressent aussi au Docteur **LEULMI Sara** et Docteur **MEZHOUD Kenza Assia**, de l'Université Frères Mentouri-Constantine 1, pour l'intérêt qu'elles ont manifesté à l'égard de mes travaux et aussi pour l'honneur qu'elles me font en prenant la charge de rapporteur et en siégeant à ce jury.

Je remercie également Docteur **TALHI Hamida**, pour sa disponibilité et pour l'honneur qu'elle me fait en acceptant d'être rapporteur de cette thèse et membre du jury.

# Dédicace

Je tiens, tout particulièrement, à témoigner une vive reconnaissance à mes très chers parents pour leur confiance inconditionnelle, leur présence à tous les instants, à m'encourager, à m'écouter à être patients dans mes exigences et parfois mes caprices.

Merci à *Maman* pour son soutien indéfectible, pour avoir traversé les nuits blanches, le stress, les doutes de la fin de thèse, sans fléchir. Nous allons pouvoir profiter ensemble de ce que la concrétisation de cette étape de ma vie, représente pour notre avenir.

Merci à *Papa* pour son enthousiasme et le soutien quotidien qu'il m'apporte, sans lequel je ne serais jamais devenue la personne que je suis aujourd'hui.

Merci à ma Tante bien-aimée *Khalissa*, mon exemple de persévérance, pour son affection, sa motivation, son implication, sa disponibilité à plein temps, ses encouragements et son assistance psychologique et scientifique sans jamais me lâcher la main.

Merci à mon Oncle bien-aimé *Rachid* qui ne cesse de m'encourager et apaiser mes tensions et mes stress par ses précieux conseils et m'enrichir de son expérience.

Je remercie ma Sœur *Selsabil*, pour sa sympathie et qui ne cesse de me porter réconfort et courage.

Je remercie mon Frère *Med Seddik*, notre cadet, qui malgré son âge m'apporte soutien et bonne humeur.

Merci à mon très cher Oncle *Murad*, un érudit des Mathématiques pour son soutien infailible, bien qu'il soit très loin de nous.

Merci à *Zineddine* et *Naila*, les deux petits anges de notre demeure qui apporte beaucoup de gaieté et de joie à mon quotidien.

Je remercie mes amis, Oh ! combien nombreux, je leurs témoigne toute ma reconnaissance. Merci à eux pour leur sympathie et leur bonne humeur. Je cite mon très cher ami *Hatem* ainsi que sa femme *Amina* d'avoir cru en moi, pour m'avoir encouragé à suivre cette voie en me souhaitant d'être heureuse ainsi.

Enfin je renouvelle mes chaleureux remerciements, mon respect et ma gratitude à ceux qui m'ont accordé du temps et m'ont témoigné un soutien constant dans ce long travail de recherche.

*L'avantage de la Science*

*Entre deux Bourgeois d'une ville  
S'émut jadis un différend :  
L'un était pauvre, mais habile ;  
L'autre riche, mais ignorant.  
Celui-ci sur son concurrent  
Voulait emporter l'avantage,  
Prétendait que tout homme sage  
Était tenu de l'honorer.  
C'était tout homme sot ; car pourquoi révéler  
Des biens dépourvus de mérite ?  
La raison m'en semble petite.  
« Mon ami, disait-il souvent au savant  
Vous vous croyez considérable ;  
Mais, dites-moi, tenez-vous table ?  
Que sert à vos pareils de lire incessamment ?  
Ils sont toujours logés à la troisième chambre,  
Vêtus au mois de juin comme au mois de décembre,  
Ayant pour tout laquais leur ombre seulement.*

*La République a bien affaire  
Des gens qui ne dépensent rien !  
Je ne sais d'homme nécessaire  
Que celui dont le luxe épand beaucoup de biens.  
Nous en usons, Dieu sait ! notre plaisir occupe  
L'artisan, le vendeur, celui qui fait sa jupe,  
Et celle qui la porte, et vous, qui dédiez  
A Messieurs les gens de la finance  
Des méchants livres bien payés »  
Ces mots remplis d'impertinence  
Eurent le sort qu'ils méritaient.  
L'homme lettré se tut, il avait trop à dire.  
La guerre le vengea bien mieux qu'une satire.  
Mars détruisit le lieu que nos gens habitaient : L'un  
et l'autre quitta sa ville.  
L'ignorant resta sans asile :  
Il reçut partout des mépris ;  
L'autre reçut partout quelque faveur nouvelle :  
Cela décida leur querelle.  
Laissez dire les sots : le savoir à son prix.*

# RÉSUMÉ

Ce travail de recherche concerne d'abord l'étude de nouvelles distributions statistiques, très flexibles, capables de décrire une large variété de données, appelées la nouvelles *Weibull-Weibull* (*NWW*) et la nouvelle *Weibull-Rayleigh* (*NWR*). Basés sur ces distributions, des modèles à durée de vie accélérée (*AFT*) ont été développés. Ces modèles *AFT-NWW* et *AFT-NWR* dépendent de variables dites explicatives qui représentent les causes ayant un impact sur la variable d'intérêt. Après avoir calculé les estimateurs du maximum de vraisemblance et les matrices d'information, des statistiques de tests permettant d'ajuster des données à ces modèles ont été construites aussi bien que dans le cas complet que dans le cas de censure à droite. Une étude par simulation massive a confirmé la faisabilité de ces tests de manière efficace. L'intérêt de ces nouvelles distributions a été illustré par des applications dans différents domaines d'étude tels que l'économie, la fiabilité et le domaine médical.

**Mots-Clés :** Analyse de survie, Données censurées, la Nouvelle famille Weibull-G, Modèle à temps de vie accéléré, Distribution Rayleigh, Distribution Weibull, Modèle N-Weibull-Rayleigh, Modèle N-Weibull-Weibull, Tests d'ajustement du chi-deux modifié.

# ABSTRACT

This work first concerns the study of new statistical distributions, very flexible, capable of describing a wide variety of data and called the New Weibull-Weibull (*NWW*) and the New Weibull-Rayleigh (*NWR*). Based on these distributions, accelerated lifetime (**AFT**) models have been developed. These *AFT-NWW* and *AFT-NWR* models depend on so-called explanatory variables which represent the causes having an impact on the variable of interest. After calculating the maximum likelihood estimators and the information matrices, goodness-of-fit test statistics were constructed for these models in both of complete and right censored data cases. A massive simulation study has confirmed the feasibility of these tests in an effective manner. The interest of these new distributions has been illustrated by applications in different fields of study such as economy, reliability and the medical field.

**Keywords :** Accelerated failure time model, Censored data, Chi-squared type Goodness-of-fit test, New-Weibull-Rayleigh model, New-Weibull-Weibull model, family, Rayleigh distribution, Survival analysis, The New-Weibull-G, Weibull distribution.

# ملخص

يتعلق هذا العمل البحثي أولاً بدراسة التوزيعات الإحصائية الجديدة، المرنة جداً، والقادرة على وصف مجموعة متنوعة من

البيانات (N-Weibull-Rayleigh (NWR و N-Weibull-Weibull (NWW).

بناءً على هذه النماذج تم تطوير نماذج العمر المتسارع (فترات حياة متسارعة) ، نقوم بطرح التوزيعات السابقة. يعتمد هذان النموذجان على ما يسمى بالمتغيرات التوضيحية التي تمثل الأسباب التي لها تأثير على متغير الاهتمام. بعد حساب تقديرات الاحتمالية القصوى ومصفوفات المعلومات، تم إنشاء إحصائيات اختبار ملاءمة لهذه النماذج في كل من حالات البيانات الكاملة والصحيحة الخاضعة للرقابة. أكدت دراسة محاكاة ضخمة جدوى هذه الاختبارات بطريقة فعالة. تم توضيح الاهتمام بهذه التوزيعات الجديدة من خلال التطبيقات في مجالات الدراسة المختلفة مثل الاقتصاد والموثوقية والمجال

الطبي.

**كلمات المفتاحية:** نموذج وقت الاخفاق المتسارع ، اختبار الضبط المعدل كا، اختبار نيكولين راو روبسون، البيانات

المراقبة، تحليل البقاء على قيد الحياة.

# Table des matières

<b>Introduction Générale</b>	<b>1</b>
<b>1 Modèles Généralisés</b>	<b>6</b>
1 Méthodes de généralisation des distributions statistiques . . . . .	6
1.1 Méthode de Pearson . . . . .	6
1.2 Méthode de transformation de Johnson . . . . .	7
1.3 Méthode de Tukey . . . . .	9
1.4 Méthode d'Eugene . . . . .	10
1.5 La famille G-Kumaraswamy . . . . .	11
1.6 La famille G-Zografos-Balakrishnan . . . . .	12
2 La nouvelle généralisation de la distribution Weibull . . . . .	13
2.1 La famille T-X . . . . .	13
2.2 La famille Weibull-X . . . . .	14
2.3 La famille Weibull-G . . . . .	15
2.4 La Nouvelle famille Weibull-G ( <i>NWG</i> ) . . . . .	15
3 Tests d'ajustements . . . . .	16
3.1 Le test de Nikulin-Rao-Robson . . . . .	18
3.2 Le test de chi-deux modifié de Bagdonavicius et Nikulin . . . . .	20
<b>2 Le modèle N-Weibull-Weibull (<i>NWW</i>)</b>	<b>27</b>

TABLE DES MATIÈRES

---

1	Présentation du modèle NWW . . . . .	28
2	Estimation . . . . .	30
2.1	Estimation des paramètres inconnus dans le cas complet . . . . .	30
2.2	La matrice d'information de Fisher estimée dans le cas complet . . . . .	32
2.3	Estimation des paramètres inconnus dans le cas censuré . . . . .	34
2.4	La matrice d'information de Fisher estimée dans le cas censuré . . . . .	36
2.5	Simulation numérique . . . . .	38
3	Test d'ajustement pour le modèle NWW . . . . .	42
3.1	Critère du test d'ajustement dans le cas complet . . . . .	42
3.1.1	Les éléments de la statistique du test . . . . .	43
3.1.2	Simulation des valeurs du critère de test . . . . .	45
3.2	Critère du test d'ajustement dans le cas censuré . . . . .	46
3.2.1	Les éléments de la statistique du test . . . . .	47
3.2.2	Simulation des valeurs du critère de test . . . . .	49
4	Tests classiques . . . . .	49
5	Applications aux données réelles . . . . .	52
5.1	Données économiques . . . . .	52
5.2	Données de fiabilité . . . . .	55
<b>3</b>	<b>Le modèle N-Weibull-Rayleigh (NWR)</b>	<b>59</b>
1	Présentation du modèle <i>NWR</i> . . . . .	60
2	Estimation . . . . .	62
2.1	Estimation des paramètres inconnus dans le cas complet . . . . .	62
2.2	La matrice d'information de Fisher estimée dans le cas complet . . . . .	64
2.3	Estimation des paramètres inconnus dans le cas censuré . . . . .	64
2.4	La matrice d'information de Fisher estimée dans le cas censuré . . . . .	66
2.5	Simulation numérique . . . . .	68
3	Test d'ajustement pour le modèle <i>NWR</i> . . . . .	71
3.1	Critère du test d'ajustement dans le cas complet . . . . .	71
3.1.1	Les éléments de la statistique du test . . . . .	72
3.1.2	Simulation des valeurs du critère de test . . . . .	73

3.2	Critère du test d'ajustement dans le cas censuré . . . . .	74
3.2.1	Les éléments de la statistique du test . . . . .	75
3.2.2	Simulation des valeurs du critère de test . . . . .	76
4	Tests classiques . . . . .	76
5	Applications aux données réelles . . . . .	78
5.1	Données sismiques . . . . .	78
5.2	Données épidémiologiques . . . . .	81
5.3	Données médicales . . . . .	84
<b>4</b>	<b>Modèle à temps de vie accéléré (AFT)</b>	<b>88</b>
1	Construction du modèle AFT-NWW . . . . .	91
1.1	Estimation des paramètres inconnus . . . . .	92
1.2	La matrice d'information de Fisher . . . . .	93
1.3	Test d'ajustement pour le modèle AFT-NWW . . . . .	96
1.4	Les éléments de la statistique du test . . . . .	97
1.5	Simulation numérique . . . . .	99
1.5.1	Les estimateurs du maximum de vraisemblance . . . . .	99
1.5.2	La statistique du test . . . . .	101
1.6	Application aux données réelles . . . . .	103
2	Construction du modèle AFT-NWR . . . . .	107
2.1	Estimation des paramètres inconnus . . . . .	108
2.2	La Matrice d'information de Fisher . . . . .	109
2.3	Test d'ajustement pour le modèle AFT-NWR . . . . .	111
2.4	Les éléments de la statistique du test . . . . .	112
2.5	Simulation numérique . . . . .	114
2.5.1	Les estimateurs du maximum de vraisemblance . . . . .	114
2.5.2	La statistique du test . . . . .	116
2.6	Application aux données réelles . . . . .	118
	<b>Conclusion</b>	<b>121</b>
	<b>Bibliographie</b>	<b>123</b>

# Liste des Figures

2.1	pdf du modèle $NWW$ vs. $WW$ . . . . .	29
2.2	hrf du modèle $NWW$ vs. $WW$ . . . . .	30
2.3	$\sqrt{n}$ -convergence des EMV de la distribution $NWW$ . . . . .	40
2.4	$\sqrt{n}$ -convergence des EMV pour les données censurées de la distribution $NWW$ . . . . .	41
2.5	pdf, cdf, QQ et PP plot de l'inflation annuelle en Algérie (1961 – 2011). . . . .	53
2.6	pdf, cdf, QQ et PP plot du temps de défaillance de 40 moteurs diesel. . . . .	56
3.1	pdf du modèle $NWR$ vs. $WR$ . . . . .	61
3.2	hrf du modèle $NWR$ vs. $WR$ . . . . .	62
3.3	$\sqrt{n}$ -convergence des EMV de la distribution $NWR$ . . . . .	69
3.4	$\sqrt{n}$ -convergence des EMV pour les données censurées de la distribution $NWR$ . . . . .	70
3.5	pdf, cdf, QQ et PP plot de la profondeur du tremblement de terre en Turquie-Syrie. . . . .	79
3.6	pdf, cdf, QQ-plot et PP-plot du taux de mortalité chez les consommateurs de drogues. . . . .	82
3.7	pdf, cdf, QQ et PP plot du temps de survie de 100 patients. . . . .	85
4.1	$\sqrt{n}$ -convergence des EMV de la distribution $AFT - NWW$ . . . . .	101

---

4.2	Histogrammes de la distribution de $Y^2$ et la distribution du chi-deux avec le degré de liberté correspondant pour la distribution $AFT - NWW$ . . . . .	103
4.3	PP et QQ plot de l'indice de masse corporelle de 100 personnes. . . . .	104
4.4	$\sqrt{n}$ -convergence des EMV de la distribution $AFT - NWR$ . . . . .	116
4.5	Histogrammes de la distribution de $Y^2$ et la distribution du chi-deux avec degré de liberté correspondant pour la distribution $AFT - NWR$ . . . . .	118
4.6	PP et QQ plot du poids de 18 tumeurs exposés à différentes doses de radioactivité. . . . .	119

# Listes des Tableaux

2.1	Les valeurs moyennes des EMV des paramètres et leurs MSE pour les données complètes du modèle <i>NWW</i> . . . . .	39
2.2	Les valeurs moyennes des EMV des paramètres et leurs MSE pour les données censurées du modèle <i>NWW</i> . . . . .	41
2.3	Comparaison des risques théoriques et des risques empiriques pour les données complètes du modèle <i>NWW</i> . . . . .	46
2.4	Comparaison des risques théoriques et des risques empiriques pour les données censurées du modèle <i>NWW</i> . . . . .	49
2.5	Valeurs des critères d'information <i>AIC/BIC/CAIC/HQIC/KS</i> pour les données simulées. . . . .	51
2.6	Données d'inflation annuelle en Algérie (1961 – 2011). . . . .	52
2.7	Valeurs des critères d'information <i>AIC, BIC, CAIC, HQIC</i> et <i>KS</i> . . . . .	54
2.8	Valeurs de $\hat{a}_j, \hat{v}_j$ et $p_j$ . . . . .	54
2.9	Le temps de défaillances de 41 moteurs diesel. . . . .	55
2.10	Valeurs des critères d'information <i>AIC, BIC, CAIC, HQIC</i> et <i>KS</i> . . . . .	57
2.11	Valeurs de $\hat{a}_j, \hat{U}_j, e_j, \hat{C}_{1j}, \hat{C}_{2j}, \hat{C}_{3j}$ et $\hat{C}_{4j}$ . . . . .	57
3.1	Les valeurs moyennes des EMV des paramètres et leurs MSE pour les données complètes du modèle <i>NWR</i> . . . . .	68

3.2	Les valeurs moyennes des EMV des paramètres et leurs MSE pour les données censurées du modèle <i>NWR</i> . . . . .	70
3.3	Comparaison des risques théoriques et des risques empiriques pour les données complètes du modèle <i>NWR</i> . . . . .	73
3.4	Comparaison des risques théoriques et des risques empiriques pour les données censurées du modèle <i>NWR</i> . . . . .	76
3.5	Valeurs des critères d'information <i>AIC/BIC/CAIC/HQIC/KS</i> pour les données simulées. . . . .	77
3.6	La profondeur du tremblement de terre en Turquie-Syrie. . . . .	78
3.7	Valeurs des critères d'information <i>AIC, BIC, CAIC, HQIC</i> et <i>KS</i> . . . . .	80
3.8	Valeurs de $\hat{a}_j, \hat{v}_j$ et $p_j$ . . . . .	80
3.9	Le taux de mortalité chez les consommateurs de drogues injectables. . . . .	81
3.10	Valeurs des critères d'information <i>AIC, BIC, CAIC, HQIC</i> et <i>KS</i> . . . . .	83
3.11	Valeurs de $\hat{a}_j, \hat{v}_j$ et $p_j$ . . . . .	83
3.12	Temps de survie de 100 patients infectés par le VIH. . . . .	84
3.13	Valeurs des critères d'information <i>AIC, BIC, CAIC, HQIC</i> et <i>KS</i> . . . . .	86
3.14	Valeurs de $\hat{a}_j, \hat{U}_j, e_j, \hat{C}_{1j}, \hat{C}_{2j}, \hat{C}_{3j}$ . . . . .	86
4.1	Les valeurs moyennes des EMV des paramètres et leurs MSE pour le modèle <i>AFT – NWW</i> . . . . .	100
4.2	Comparaison des risques théoriques et des risques empiriques du modèle <i>AFT – NWW</i> . . . . .	102
4.3	L'indice de masse corporelle mesuré à différentes valeurs de la tension artérielle systolique. . . . .	104
4.4	Valeurs de $\hat{a}_j, \hat{v}_j$ et $p_j$ . . . . .	105
4.5	Les valeurs moyennes des EMV des paramètres et leurs MSE du modèle <i>AFT – NWR</i> . . . . .	115
4.6	Comparaison des risques théoriques et des risques empiriques de la distribution <i>AFT – NWR</i> model. . . . .	117
4.7	Le poids de 18 tumeurs exposés à différentes doses de radioactivité. . . . .	119
4.8	Valeurs de $\hat{a}_j, \hat{v}_j$ et $p_j$ . . . . .	120

# Introduction Générale

L'analyse de survie est une méthode statistique utilisée pour analyser le délai de la survenue au cours du temps d'un évènement d'intérêt, tel qu'un échec, un décès, une rechute ou une rémission. L'évènement dans l'analyse de survie est souvent appelé un « évènement d'intérêt », tandis que le temps nécessaire pour que l'évènement se produise est appelé « temps de survie ». Pour mener à bien une analyse de survie sur une population précise, il est nécessaire de choisir le bon modèle statistique qui décrira de la meilleure manière la relation entre les variables d'un ensemble de données, faire des prédictions et tester des hypothèses. Depuis le *XXe* siècle les chercheurs ne cessent de montrer de l'intérêt aux méthodes statistiques dans la modélisation des temps de survie et de défaillance des évènements précis.

La première méthode actuarielle est apparue en 1912, elle a été utilisée dans le domaine médical pour la première fois en 1950, la seconde méthode, dite Kaplan-Meier, est apparue en 1958, ces méthodes modélisent des observations complètes, ce qui n'est pas toujours le cas pour les données de survie. L'analyse de survie prend en compte le fait que l'évènement d'intérêt de certains individus ou sujets peut ne pas survivre jusqu'à la fin de l'analyse, ce qui entraîne des données censurées ; c'est à dire l'apparition au cours du temps d'un évènement avec des observations « incomplètes » constituées par

les sujets chez qui l'évènement n'est pas encore survenu à la fin de l'analyse.

Depuis le *XXI* siècle, les données exploitées deviennent plus complexes et difficiles à manipuler, il arrive souvent que ces données dépendent de différents facteurs appelés stress qui influencent la durée de vie, par exemple, la température, le poids, la dose de médicaments, etc. Afin de modéliser et analyser ces données, les chercheurs ont eu recours aux modèles à temps de vie accéléré AFT (Accelerated Failure Time). Pour ces modèles, les fonctions d'intérêt sont des fonctions de variables dites variables explicatives représentant les différents facteurs de stress, en d'autres termes, les temps de survie des individus ou des produits sont accélérés ou ralentis par un facteur déterminé représenté par une variable explicative. Les modèles AFT estiment directement l'effet des variables explicatives sur le temps de survie du produit, ainsi, nous pouvons fournir une estimation de la fiabilité du matériel dans des délais courts permettant d'optimiser le temps, la performance et la planification de la maintenance de celui-ci.

Les modèles AFT sont couramment utilisés dans divers domaines tels que la recherche médicale et l'épidémiologie notamment dans le suivi thérapeutique du dosage des médicaments (TDM) comme réponse au traitement (temps de survie) ou une dysfonction d'organe d'un patient en analyse de survie, dans les sciences sociales où les différents facteurs (la situation financière des parents, l'organisme scolaire, les compétences professionnelles, etc...) pouvant influencer sur le rendement scolaire, dans l'ingénierie pour optimiser la fiabilité des systèmes, en finance et dans les études commerciales pour analyser les effets comportementaux des clients. Les modèles AFT permettent aux chercheurs de comprendre les facteurs influençant le délai avant un évènement et de faire des prédictions sur les résultats futurs en fonction des probabilités de survie.

Ces modèles sont proposés comme une alternative aux modèles à risques proportionnels de Cox (1972) connus sous le nom de modèles PH (proportional hazards) où la condition de proportionalité des taux de hasard est exigée. Différents modèles AFT ont été proposés dans la littérature statistique tels que AFT-Weibull, AFT-exponentiel, AFT-

log-logistique, AFT-log-normale ( Bagdonavicius et Nikulin, 2002), AFT-Inverse Weibull généralisée (Goual et Seddik-Ameur, 2014), AFT-exponential-exponentié -généralisé (Aidi et Seddik-Ameur, 2016), AFT-extension de Weibull (Treidi et Seddik-Ameur, 2017).

Prenant en considération la révolution industrielle et le développement technologique, les données réelles peuvent souvent être complexes et présentent divers défis qui doivent être relevés au cours du processus de la modélisation statistique, les distributions classiques existantes ne sont pas adaptées pour modéliser toutes les observations. Il a été crucial d'introduire de nouveaux générateurs de distributions afin de pouvoir traiter et analyser toutes les données réelles disponibles. La construction de nouveaux modèles statistiques flexibles aptes à générer une grande variété de formes du taux de risque (taux de défaillance) fréquemment observées est devenue plus que nécessaire.

Durant les deux dernières décennies, les chercheurs ont proposé diverses techniques de généralisation des distributions classiques donnant ainsi différentes formes pour le taux de hasard aussi bien croissante, décroissante, unimodale, en baignoire, en forme J et S. Parmi ces techniques, plusieurs sont basées sur la distribution de Weibull telles que la famille bêta Weibull-G [Cordeiro et al. (2012)], la Weibull-X [Alzaatrah et al. (2013)], la Weibull-X exponentiée [Alzaghal et al. (2013)] et la Weibull-G [Bourguignon et al. (2014)].

Dans la présente thèse, nous nous sommes intéressées à un nouveau générateur de modèles, basé sur la distribution de Weibull introduit récemment par *Tahir et al.* (2016). Cette nouvelle famille est appelée la Nouvelle famille Weibull-G (New-Weibull-G family, *NWG*).

Nous choisissons la distribution de Weibull et celle de Rayleigh comme distributions de base de cette famille pour construire deux nouveaux modèles statistiques appelés le modèle New-Weibull-Weibull (*NWW*) et le modèle New-Weibull Rayleigh (*NWR*). Ces modèles sont très intéressants dans la modélisation de données provenant de tous

les domaines d'application que ce soit en finance, en fiabilité, en analyse de survie ou autres.

Étant donné que dans les études de fiabilité ou des durées de vie, le temps de l'occurrence de l'évènement peut être très long, on a jugé utile de développer des modèles à temps de vie accéléré correspondants à ceux proposés dans ce travail,  $AFT - NWW$  et  $AFT - NWR$ .

Après la description de ces nouvelles distributions  $NWW$ ,  $NWR$ ,  $AFT - NWW$  et  $AFT - NWR$ , nous déterminons les estimateurs des paramètres inconnus. Pour valider le choix de ces modèles, nous utilisons d'abord les critères classiques de sélection des modèles tels que le PP-plot, le QQ plot, l'AIC, le BIC, CAIC, HQIC et KS; ensuite nous nous proposons de construire des statistiques de tests du type du chi-deux modifié relatifs à chacun de ces modèles dans le cas de données complètes et le cas de données censurées à droite. On utilise l'approche proposée par Nikulin (1973) et Rao et Robson (1974) dans le cas de données complètes et la modification de cette statistique [Nikulin et Bagdonavicius, (2011)] pour les données censurées à droite.

Ces critères de tests sont basés sur la méthode d'estimation du maximum de vraisemblance sur les données initiales utilisant ainsi toute l'information apportée par l'échantillon. Pour confirmer les résultats obtenus, on a procédé à la simulation de dizaines de milliers d'échantillons de plusieurs tailles et différentes valeurs des paramètres. A la fin, on a utilisé les statistiques de tests développées dans ce travail pour ajuster aux modèles étudiés plusieurs jeux de données réelles provenant des domaines médical, industriel, économique, social et sismique montrant ainsi la flexibilité et l'applicabilité de ces tests.

La thèse est structurée de la manière suivante :

Dans le premier chapitre, les différentes méthodes de généralisations de la distribution de Weibull ont été détaillées. Une brève introduction sur la théorie des tests d'ajustement du type du chi-deux modifié dans le cas de données complètes et le cas de don-

nées censurées à droite est exposée.

Dans le chapitre deux, on introduit le nouveau modèle N-Weibull-Weibull (*NWW*). Une analyse statistique complète a été réalisée. La statistique de test d'adéquation à ce nouveau modèle a été calculée pour le cas de données complètes et le cas de données censurées à droite. Des applications sur des données réelles ont été présentées afin de montrer la flexibilité de cette nouvelle distribution *NWW*.

Quant au troisième chapitre, il est consacré au modèle N-Weibull-Rayleigh (*NWR*). Après avoir calculé les estimateurs inconnus du modèle par la méthode du maximum de vraisemblance, la statistique du test du chi-deux modifiée introduite dans le premier chapitre a été adaptée au modèle proposé et la forme explicite de tous les éléments du critère de test été obtenue pour le cas complet et le cas de censure à droite. Les résultats théoriques ont été appliqués à des jeux de données réelles.

Dans le dernier chapitre, après l'introduction de la notion des modèles à temps de vie accéléré, on construit les modèles *AFT – NWW* et *AFT – NWR* dont les distributions de base sont respectivement la N-Weibull-Weibull et la N-Weibull-Rayleigh. En utilisant les estimateurs du maximum de vraisemblance des paramètres inconnus de la distribution de base et ceux de la régression, on construit des tests d'ajustement du type du chi-deux modifié pour valider ces deux modèles. D'importantes simulations numériques ont été conduites pour confirmer les résultats théoriques obtenus. Des applications à des jeux de données réelles provenant de différents domaines ont illustré l'applicabilité des modèles étudiés et la souplesse des tests proposés dans ce travail de recherche.

Enfin, nous espérons, au travers de ce travail, apporter une « brique » conceptuelle à un cadre de recherche sur l'importance et la performance des distributions statistiques dans la traduction de différents types de données rencontrées dans le monde réel.

# Modèles Généralisés

## 1 Méthodes de généralisation des distributions statistiques

Les distributions statistiques sont couramment utilisées pour décrire des phénomènes réels. Récemment, de nombreuses classes de distributions généralisées ont été développées pour modéliser diverses observations.

Il existe des méthodes bien connues pour générer des distributions continues univariées. Ces méthodes comprennent la méthode de l'équation différentielle de Pearson (1895), la méthode de transformation de Johnson (1949), la méthode des fonctions quantiles de Tukey (1960) et d'autres méthodes décrites ci-dessous.

### 1.1 Méthode de Pearson

Vers la fin du dix-neuvième siècle, *K. Pearson* (1895) a voulu construire un nouveau système de courbes de fréquences qui représentera une plus grande variété de distributions que celles pour lesquelles une courbe normale suffirait.

Le système de distributions continues de *Pearson*, développé en (1895) est un système pour lequel toute fonction de densité de probabilité  $f(x)$  satisfait une équation différentielle de la forme :

$$\frac{1}{f(x)} \frac{df(x)}{dx} = \frac{a+x}{b_0 + b_1x + b_2x^2},$$

Où  $a$ ,  $b_0$ ,  $b_1$  et  $b_2$  représentent les paramètres. Dépendant de ceux-ci, les différentes formes de la distribution  $f(x)$  ont été classées par *Pearson* dans un certain nombre de types.

En 1942, *Burr* a présenté un système de distributions continues pouvant prendre une grande variété de formes. Le système de distributions satisfait l'équation différentielle suivante :

$$dF = F(1 - F)g(x)dx,$$

Où  $0 \leq F \leq 1$  et  $g(x)$  est une fonction non négative sur  $x$ . *Burr* a donné 12 solutions de l'équation différentielle et cela correspond aux choix de  $g(x)$ , voir *Fry* (1993) et *Johnson et al.* (1994) pour une liste de distributions de *Burr* types *I - XII*.

### 1.2 Méthode de transformation de Johnson

Pour obtenir un système de courbes analogue aux systèmes de *Pearson* où  $f(x)$  doit être définie et de préférence sous une forme simple, *Johnson* (1949) a proposé un système pour générer des distributions en utilisant la transformation de normalisation avec la forme générale :

$$Z = \gamma + \delta f\left(\frac{x - \xi}{\lambda}\right),$$

où  $f(\cdot)$  est la fonction de transformation,  $Z$  est une variable aléatoire normale centrée et réduite,  $\gamma$  et  $\delta$  sont les paramètres de formes,  $\lambda$  représente le paramètre d'échelle et  $\xi$  c'est le paramètre de position. *Johnson* a supposé que  $\delta$  et  $\lambda$  sont positifs. Il a proposé trois fonctions de transformation et a défini la famille lognormale (*SL*), le système de distribution borné (*SB*) et le système de distribution non borné (*SU*). Ces familles de distributions couvrent de nombreuses distributions couramment utilisées telles que normale, log-normale, gamma, bêta, exponentielle et d'autres.

Les tests de signification ne sont pas vérifiés dans les systèmes développés par *Johnson*.

Les systèmes proposés par *Johnson SL, SB, SU* se combinent avec la courbe normale pour donner une variété de formes de courbes aussi large que celles fournies par les systèmes de courbes de fréquences de *Pearson*.

Le système log-normale (*SL*) s'est avéré utile dans un certain nombre d'applications. Il modélise les problèmes de dosage-mortalité (*Gaddum, 1945*), de la graduation des données économiques (*Gibrat, 1931; Frechet, 1945*) et dans l'agriculture (*Cochran, 1938*).

Le système de distribution borné (*SB*) est basé sur la transformation la plus simple, mais le système de distribution non borné (*SU*) a les expressions les plus simples pour déterminer les moments. L'ajustement du système de distribution non borné (*SU*) est plus simple que celui du système de distribution borné (*SB*) sauf lorsque les limites de variation de ce dernier sont définitivement connues.

Tous les moments de toutes les courbes dans les systèmes proposés par *Johnson SL, SB* et *SU* sont finis. En comparaison, on sait que les moments les plus élevés de certaines des courbes de *Pearson* peuvent être infinis.

### 1.3 Méthode de Tukey

Tukey (1960) a proposé la distribution lambda, qui a été généralisée par Ramberg et Schmeiser [(1972), (1974)] et Ramberg *et al.*, (1979) comme une distribution nommée lambda généralisée (*GLD*).

Cette famille de distribution est définie en termes de fonction quantile :

$$Q(y) = Q(y; \lambda_1, \lambda_2, \lambda_3, \lambda_4) = \lambda_1 + \frac{y^{\lambda_3} - (1-y)^{\lambda_4}}{\lambda_2}, 0 \leq y \leq 1,$$

Les paramètres  $\lambda_1$  et  $\lambda_2$  sont, respectivement, les paramètres de position et d'échelle, tandis que  $\lambda_3$  et  $\lambda_4$  déterminent les coefficients de l'asymétrie et de l'aplatissement. La fonction de densité de probabilité correspondante est donnée par :

$$f(x) = \frac{\lambda_2}{\lambda_3 y^{\lambda_3-1} + \lambda_4 (1-y)^{\lambda_4-1}}, x = Q(y),$$

L'existence d'une fonction de densité de probabilité valide exige que  $\lambda_3 y^{\lambda_3-1} + \lambda_4 (1-y)^{\lambda_4-1}$  ait le même signe pour tout  $y$  dans  $[0, 1]$  et  $\lambda_2$  prend le même signe.

Freimer *et al.* (1988) ont discuté la similitude et les différences entre le système de Pearson et la *GLD*. Ils ont souligné que la famille de Pearson n'inclut pas la distribution logistique, tandis que *GLD* ne couvre pas toutes les valeurs d'asymétrie et d'aplatissement.

Une extension de *GLD* a été proposée par Karian et Dudewicz (2000) qui comprend à la fois *GLD* et la distribution bêta généralisée définie par :

$$f(x) = \left\{ \begin{array}{l} \frac{(x-\beta_1)^{\beta_3} (\beta_1+\beta_2-x)^{\beta_4}}{B(\beta_3+1, \beta_4+1) \beta_2^{(\beta_3+\beta_4+1)}}, \text{ si } \beta_1 \leq x \leq \beta_1 + \beta_2 \\ 0, \quad \text{sinon} \end{array} \right\}.$$

Où  $B(.,.)$  est la fonction bêta complète.

## 1.4 Méthode d'Eugene

La distribution Lambda généralisée (*GLD*), initialement proposée par Ramberg et Schmeiser (1974) appelée distribution "RS", est une généralisation à quatre paramètres de la famille Lambda de Tukey (Hastings *et al.*, 1947) qui s'est révélée utile dans un certain nombre d'applications différentes. Puisqu'elle peut prendre une grande variété de formes, la *GLD* offre aux gestionnaires de risques une grande flexibilité dans la modélisation d'un large éventail de données financières. En raison de sa polyvalence, l'obtention des paramètres appropriés pour la *GLD* peut être un problème difficile.

Eugene *et al.* (2002) ont utilisé la distribution bêta en tant que générateur pour développer la famille appelée distribution de bêta-généralisée. La fonction de répartition de la variable aléatoire  $X$  de cette famille est définie par :

$$G(x) = \int_0^{F(x)} b(t) dt,$$

où,  $b(t)$  est la fonction de densité de probabilité de la variable aléatoire bêta et  $F(x)$  est la fonction de répartition d'une variable aléatoire quelconque.

La fonction de densité correspondante à la distribution bêta-généralisée est donnée par :

$$g(x) = \frac{1}{B(\alpha, \beta)} f(x) F^{\alpha-1}(x) (1 - F(x))^{\beta-1}.$$

Cette famille de distributions est une généralisation des distributions de la statistique d'ordre pour la variable aléatoire  $X$  avec une fonction de répartition  $F(x)$  indiquée par Eugene *et al.*, (2002) et Jones (2004).

Depuis l'article d'Eugène *et al.*, (2002), de nombreuses distributions bêta-généralisée ont été étudiées dans la littérature incluant la distribution bêta-Gumbel (Nadarajah et Kotz, 2004), la distribution bêta-Exponentielle (Nadarajah et Kotz, 2005), la distribution bêta-Weibull (Famoye *et al.*, 2004), la distribution bêta-Gamma (Kong *et al.*, 2007), la distribution bêta-Pareto (Akinsete *et al.*, 2008) et autres.

### 1.5 La famille G-Kumaraswamy

Jones (2009) et Cordeiro et de Castro (2011) ont prolongé la famille de distributions bêta-généralisée en remplaçant la distribution bêta de la distribution bêta-généralisée par la distribution de Kumaraswamy (1980) :

$$f(x) = \begin{cases} \alpha\beta x^{\alpha-1}(1-x^\alpha)^{\beta-1} & ; x \in [0, 1], \alpha > 0, \beta > 0 \\ 0 \text{ sinon} & \end{cases}$$

Alors, la fonction de densité de la distribution Kumaraswamy généralisée (Cordeiro et de castro, 2011) est donnée par :

$$g(x) = \alpha\beta f(x) F^{\alpha-1}(x) (1 - F^\alpha(x))^{\beta-1},$$

Plusieurs distributions généralisées de la Kumaraswamy ont été étudiées dans la littérature y compris la distribution de Kumaraswamy Weibull (Cordeiro *et al.*, 2010), la distribution Kumaraswamy gamma généralisée (De Castro *et al.*, 2011), la distribution Kumaraswamy semi-normale généralisée (Cordeiro et al., 2012) et la distribution Kumaraswamy inverse Weibull généralisée (Goual et Seddik Ameer, 2016).

## 1.6 La famille G-Zografos-Balakrishnan

Pour toute fonction de répartition de base  $G(x)$ , Zografos et Balakrishnan (2009) ont défini une distribution avec fonction de densité de probabilité  $f(x)$  et fonction de répartition  $F(x)$ , données respectivement par :

$$f(x) = \frac{1}{\Gamma(a)} \{-\log[1 - G(x)]\}^{a-1} g(x),$$

$$F(x) = \frac{\gamma(a, \log[1 - G(x)])}{\Gamma(a)} = \frac{1}{\Gamma(a)} \int_0^{-\log[1 - G(x)]} t^{a-1} \exp(-t) dt, a > 0,$$

où  $g(x) = \frac{dG(x)}{x}$ ,

$$\Gamma(a) = \int_0^\infty t^{a-1} \exp(-t) dt,$$

représente la fonction gamma, et

$$\gamma(a, z) = \int_0^z t^{a-1} \exp(-t) dt,$$

est la fonction gamma incomplète.

Récemment, Alexander *et al.*, (2012) ont étudié la famille bêta- $X$  généralisée en considérant la distribution bêta généralisée du premier type introduite par McDonald (1984). La nouvelle famille est appelée famille généralisée de bêta (*GBG*).

Alzaatreh *et al.*, (2013) ont proposé une méthode générale en remplaçant la fonction de densité de probabilité de bêta de la famille *GBG* par la fonction de densité de probabilité d'une variable aléatoire continue quelconque et en appliquant une fonction  $D(F(x))$  satisfaisant certaines conditions qui permettra de définir la famille  $T - X$ .

## 2 La nouvelle généralisation de la distribution Weibull

La distribution de *Weibull* est l'un des modèles les plus populaires et les plus utilisés pour le temps de défaillance dans les tests de vie, de survie et la théorie de la fiabilité. Cependant, son inconvénient majeur est le comportement monotone de sa fonction du taux de hasard.

Dans les applications réelles, les courbes du taux de hasard empirique présentent souvent des formes non-monotones telle qu'une baignoire, baignoire inversée (unimodale) et autres. Alors, il y a un véritable besoin de chercher des généralisations ou modifications de la distribution *Weibull* pouvant fournir plus de flexibilité dans la modélisation des durées de vie.

La fonction de répartition d'une variable aléatoire continue  $X$  qui suit une distribution *Weibull* de paramètre d'échelle  $\alpha > 0$  et de paramètre de forme  $\beta > 0$ , est donnée par :

$$F_W(x) = 1 - \exp(-\alpha x^\beta), \quad x > 0$$

La fonction de densité de probabilité correspondante est :

$$f_W(x) = \alpha \beta x^{\beta-1} \exp(-\alpha x^\beta).$$

### 2.1 La famille T-X

Soit  $r(t)$  la fonction de densité de probabilité de la variable aléatoire  $T \in [a, b]$ , tel que  $\infty < a < b < \infty$ . Soit  $F(x)$  la fonction de répartition de la variable aléatoire  $X$  et la fonction de lien  $D(\cdot) : [0, 1] \rightarrow [a, b]$  qui satisfait les conditions suivantes :

*i.*  $D(\cdot)$  est différentiable et monotone non décroissante.

*ii.*  $D(0) \rightarrow a$  et  $D(1) \rightarrow b$ .

La famille de distributions  $T-X$  a été définie en 2013 par Alzaatreh *et al.*, par la fonction de répartition donnée comme suit :

$$F(x) = \int_a^{D[G(x)]} r(t) dt, \quad (*)$$

où  $G(x)$  est la fonction de répartition de la distribution de base et  $D[G(x)]$  satisfait les conditions (i) et (ii).

Si  $T \in (0, \infty)$ ,  $X$  est une variable aléatoire continue et  $D[G(x)] = -\log[1 - G(x)]$ , alors, la fonction de densité de probabilité de la famille  $T-X$  est donnée par :

$$f(x) = \frac{g(x)}{1 - G(x)} r(-\log[1 - G(x)]) = h_g(x) r[H_g(x)],$$

où  $h_g(x)$  et  $H_g(x)$  sont les fonctions du taux de hasard et du taux de hasard cumulé, respectivement associées à  $g(x)$ .

## 2.2 La famille Weibull-X

Si  $r(t) \rightsquigarrow F_W(x)$ , posons  $D[G(x)] = -\log[1 - G(x)]$  dans (\*), alors, la fonction de répartition de la famille Weibull-X de Alzaatrah *et al.*, (2013) est :

$$\begin{aligned} F(x) &= \alpha\beta \int_0^{-\log[1-G(x)]} x^{\beta-1} \exp(-\alpha x^\beta) dt \\ &= 1 - \exp\left(-\alpha \{-\log[1 - G(x)]\}^\beta\right). \end{aligned}$$

### 2.3 La famille Weibull-G

Un an plus tard, Bourguignon *et al.*, (2014) ont proposé la famille de distribution Weibull-G en remplaçant l'argument  $x$  par :  $\frac{G(x,\theta)}{\overline{G}(x,\theta)}$ , où :  $\overline{G}(x,\theta) = 1 - G(x,\theta)$  et ont défini la fonction de répartition de leur classe (pour  $\alpha > 0$  et  $\beta > 0$ ), appelée Weibull-G, par :

$$\begin{aligned} F(x, \alpha, \beta, \theta) &= \alpha\beta \int_0^{\left[\frac{G(x,\theta)}{\overline{G}(x,\theta)}\right]} t^{\beta-1} \exp(-\alpha t^\beta) dt \\ &= 1 - \exp\left(-\alpha \left[\frac{G(x,\theta)}{\overline{G}(x,\theta)}\right]^\beta\right), x \in \mathbb{R}. \end{aligned}$$

Où  $G(x,\theta)$  et  $g(x,\theta)$  sont, respectivement la fonction de répartition et la fonction de densité de probabilité de toute distribution de base qui dépend d'un vecteur de paramètre  $\theta$ .

Dans la littérature récente, quatre générateurs basés sur la distribution *Weibull* sont apparus : bêta *Weibull-G* [Cordeiro *et al.* (2012)], *Weibull-X* [Alzaatrah *et al.* (2013)], *Weibull-X* exponentié [Alzagal *et al.* (2013)] et *Weibull-G* [Bourguignon *et al.* (2014)].

### 2.4 La Nouvelle famille Weibull-G (NWG)

En 2016, Tahir *et al.* ont proposé une classe de distributions appelée la Nouvelle famille *Weibull-G* "NWG". En remplaçant  $a = 0$ ,  $r(t) \rightsquigarrow F_W(x)$  et  $D[G(X)] = -\log[G(x,\xi)]$ , de l'équation (\*), on définit la fonction de répartition de la famille *NWG* par :

$$\begin{aligned} F(x, \alpha, \beta, \xi) &= 1 - \int_0^{-\log[G(x,\xi)]} \alpha\beta t^{\beta-1} \exp(-\alpha t^\beta) dt \\ &= \exp\left(-\alpha \{-\log[G(x,\xi)]\}^\beta\right). \end{aligned}$$

Par la suite, la variable aléatoire  $X$  est notée par :  $X \rightsquigarrow NWG(\alpha, \beta, \xi)$ . Alors, la fonction de densité de probabilité de  $X$  est donnée par :

$$f(x, \alpha, \beta, \xi) = \alpha \beta \frac{g(x, \xi)}{G(x, \xi)} \{-\log [G(x, \xi)]\}^{\beta-1} \exp\left(-\alpha \{-\log [G(x, \xi)]\}^\beta\right),$$

où  $g(x, \xi)$  est la fonction de densité de probabilité parente. De plus, parfois on peut négliger la dépendance  $\xi$  avec les paramètres et on écrit  $G(x, \xi) = G(x)$  et  $g(x, \xi) = g(x)$ . La densité de probabilité sera plus traitable lorsque la fonction de répartition  $G(x)$  et la fonction de densité de probabilité  $g(x)$  ont des expressions analytiques simples.

Soit  $h(x, \xi)$  la fonction du taux de hasard de la fonction parente  $G$ . La fonction du taux de hasard  $h(x, \alpha, \beta, \xi)$  de la variable aléatoire  $X$  de la Nouvelle famille Weibull-G ( $NWG$ ) peut être obtenue à l'aide de l'équation suivante :

$$h(x, \alpha, \beta, \xi) = \frac{\alpha \beta \frac{g(x, \xi)}{G(x, \xi)} \{-\log [G(x, \xi)]\}^{\beta-1} \exp\left(-\alpha \{-\log [G(x, \xi)]\}^\beta\right)}{1 - \exp\left(-\alpha \{-\log [G(x, \xi)]\}^\beta\right)}.$$

### 3 Tests d'ajustements

Les tests de qualité d'ajustement indiquent si un échantillon aléatoire provient d'une distribution de probabilité spécifique. Cette technique statistique est fréquemment utilisée dans l'analyse de survie et la fiabilité pour la modélisation précise de la durée de vie qui nécessite de spécifier le modèle statistique adéquat. Au début du 20<sup>ème</sup> siècle, *Karl Pearson* (1900) a été le premier à lancer le test d'ajustement ( $GOF$ ) avec l'introduction de la statistique du test de chi-deux ( $X^2$ ) connue sous le nom de la statistique de *Pearson* jusqu'à ce jour. La statistique de *Pearson* nécessite que les données soient d'abord regroupées, complètes et les paramètres du modèle choisi connus. Dans l'application réelle, ce test ne peut être appliqué dans toutes les situations, particulière-

ment lorsque les paramètres du modèle statistique sont inconnus ou bien les données ne sont pas complètes. Donc pour résoudre ces problèmes, de nombreuses modifications ont été apportées au test du chi-deux de *Pearson*, telles-que *Fisher* (1928), *Cramer* (1946), *Anderson-Darling* (1952), *Chernoff* et *Lehmann* (1954), *Roy* (1956), et autres, ainsi que la proposition de tests graphiques.

Cependant, leurs utilisations présentent certains inconvénients; par exemple, selon *Fisher* (1928) la distribution limite de la statistique de *Pearson* ne suit pas la distribution de chi-deux  $\chi^2$  si les paramètres sont inconnus et doivent être estimés à partir de l'échantillon car la distribution limite dépend de la méthode d'estimation des paramètres.

*Cramer* (1948) a montré que si les paramètres inconnus du modèle ajusté sont estimés à l'aide de la méthode du minimum chi-deux ou de la méthode d'estimation du maximum de vraisemblance (EMV) basée sur des données groupées et l'estimateur est  $\sqrt{n}$ -convergent, alors, la distribution limite  $\chi_n^2$  suit un chi-deux mais le degré de liberté sera réduit du nombre de paramètres estimés.

Par la suite, en 1954, *Chernoff* et *Lehmann* ont clairement montré que la distribution limite de la statistique de *Pearson* change complètement et ne suit pas la distribution  $\chi^2$  si les paramètres de l'hypothèse nulle sont estimés à partir de l'échantillon par la méthode du maximum de vraisemblance.

Alors, il est nécessaire de trouver un meilleur test d'ajustement capable de surmonter ces obstacles, et qui fournit suffisamment d'informations de base sur l'échantillon. Notons qu'il existe un célèbre test de chi-deux modifié introduit par *Nikulin* (1973) qui est une modification de la statistique de *Pearson* (1900). Ce test du chi-deux modifié est basé sur les données initiales pour estimer les paramètres du modèle et où les bornes des intervalles de regroupement des données initiales sont calculées de manière à avoir des classes équiprobables.

La distribution limite de cette statistique de test suit la distribution du chi-deux.

Cette statistique utilise essentiellement la propriété de la distribution normale multivariée du vecteur qui repose sur les différences entre deux estimateurs des probabilités d'appartenir aux intervalles de regroupement, un estimateur basé sur la fonction de répartition empirique et un autre sur l'estimation du maximum de vraisemblance des paramètres de l'hypothèse nulle. Ensuite, en utilisant les propriétés asymptotiques de la distribution normale multivariée "best asymptotically normal" de l'estimateur du maximum de vraisemblance et de la méthode de Wald, Nikulin a complètement résolu le problème pour toute hypothèse nulle  $H_0$ .

Un an plus tard, Rao-Robson (1974) obtiennent le même résultat pour la famille de distributions exponentielles en utilisant la méthode des moments pour estimer les paramètres. Dès lors, cette statistique est appelée statistique Nikulin Rao-Robson (*NRR*), pour plus de détails voir Van der Vaart (1998), Drost (1988).

### 3.1 Le test de Nikulin-Rao-Robson

Soit  $X = (X_1, \dots, X_n)^T$  un  $n$ -échantillon, *i.e.*  $X_1, \dots, X_n$  sont des variables aléatoires *i.i.d.* Supposons qu'on a une famille de distribution  $F(\bullet, \theta)$  dans  $\mathbf{R}$ , où  $\theta = (\theta_1, \theta_2, \dots, \theta_s) \in \Theta \subset \mathbf{R}^s$ , est le vecteur des paramètres inconnus. Nous souhaitons tester l'hypothèse nulle  $H_0$  selon laquelle :

$$H_0 : \mathbf{P}_\theta(X_i \leq x) = F(x, \theta), \quad x \in \mathbf{R}^1, \theta = (\theta_1, \dots, \theta_s) \in \Theta \subset \mathbf{R}^s.$$

*Nikulin* (1973) et *Rao-Robson* (1974) ont proposé la statistique du test de chi-deux modifié pour la famille de distributions continues, appelée la statistique de *Nikulin-Rao-Robson* (*NRR*).

Basée sur les estimateurs du maximum de vraisemblance sur les données non groupées, cette statistique de test ( $Y^2$ ) peut être écrite comme la somme de la célèbre statistique de *Pearson*  $X_n^2$  et d'une forme quadratique  $Q$  :

$$\begin{aligned} Y^2 &= X_n^2 + Q, \\ Q &= \frac{1}{n} L^T(\hat{\theta}) G^{-1} L(\hat{\theta}). \end{aligned}$$

Les données initiales sont groupées en  $r$  intervalles  $I_j = ]a_{j-1}, a_j]$ , tel que  $j = \overline{1, r}$ . Les classes sont supposées équiprobables, les probabilités correspondantes sont données par :

$$p_j(\hat{\theta}_n) = \int_{a_{j-1}}^{a_j} dF(t, \hat{\theta}_n) = \frac{1}{r}, \quad j = \overline{1, r},$$

par conséquent, les limites  $a_j$  sont obtenues telles que :

$$a_j = F^{-1}\left(\frac{j}{r}\right), \quad j = \overline{1, r}.$$

On a  $v = (v_1, \dots, v_r)^T$  représente les fréquences empiriques obtenues par le regroupement des données initiales dans les intervalles  $I_j$  :

$$v_j = \text{card} \{i : x_i \in I_j, \} \quad i = 1, 2, \dots, n, \quad j = \overline{1, r}$$

Alors, les composants de  $Y^2$  sont donnés comme suit :

$$\begin{aligned}
 X_n^2 &= \sum_{j=1}^r \frac{(v_j - np_j)^2}{np_j}, \quad Q = \frac{1}{n} L^T(\hat{\theta}) G^{-1} L(\hat{\theta}), \quad \theta = (\theta_1, \theta_2, \dots, \theta_s) \\
 L(\hat{\theta}) &= [L_1(\hat{\theta}), \dots, L_s(\hat{\theta})]^T, \quad L_k(\theta) = \sum_{j=1}^r \frac{v_j}{p_j} \frac{\partial p_j(\theta)}{\partial \theta_k}, \quad k = 1, \dots, s, \quad j = 1, \dots, r, \\
 \hat{G} &= [\hat{g}_{ll'}]_{r \times r}, \quad g_{ll'} = i_{ll'} - \sum_{j=1}^r \frac{1}{p_j} \left( \frac{\partial p_j(\theta)}{\partial \theta_l} \frac{\partial p_j(\theta)}{\partial \theta_{l'}} \right), \quad \hat{i}_{ll'} = \frac{\partial^2 l(x_i, \hat{\theta})}{\partial \hat{\theta}_l \partial \hat{\theta}_{l'}},
 \end{aligned}$$

où  $I = [i_{ll'}]_{s \times s}$  représente la matrice d'information de Fisher.

Notez que cette approche a été utilisée pour l'ajustement de différentes distributions récentes, un modèle à risques concurrents (Chouia and Seddik-Ameur, 2014), la distribution de Rayleigh généralisée (Tilbi et Seddik-Ameur, 2017), l'extension de Weibull (Treidi et Seddik-Ameur, 2016) et d'autres.

### 3.2 Le test de chi-deux modifié de Bagdonavicius et Nikulin

Dans la modélisation statistique des phénomènes réels on est souvent confronté à des données incomplètes ce qui nous empêchera d'appliquer la statistique du test *NRR*. Ces obstacles ont motivé plusieurs chercheurs à proposer des techniques d'ajustement adaptées, on peut citer Habib et Thomas (1986), Hollander et Pena (1992) qui ont développé une statistique du chi-deux de type *Pearson*, *Akritas* (1988) qui a proposé une statistique du chi-deux basée sur l'idée de comparer le nombre d'échecs observé et attendu dans chaque classe, *Hjort* (1990) qui lui, a développé une statistique de type chi-deux pour tester la validité du modèle paramétrique pour les données d'histoire de vie basée sur le processus de risque cumulatif et *Kim* (1993) a également proposé un test d'ajustement du chi carré basé sur l'estimateur de la limite du produit.

*Bagdonavicius* et *Nikulin* (2011), *Bagdonavicius et al.*, (2010, 2013) ont étendu la même idée de statistique *NRR* aux données censurées. Ils ont estimé les bornes des intervalles

$a_j$  en tant que fonctions de données aléatoires basées sur l'idée de diviser l'intervalle  $[0, \tau]$  en  $k$  sous-intervalles avec des nombres de défaillance attendus égaux  $e_j$ . Le test est basé sur le vecteur  $Z = (Z_1, \dots, Z_k)^T$ , où  $Z_j = \frac{1}{\sqrt{n}}(U_j - e_j)$ ,  $j = \overline{1, k}$ , *i.e.* les différences entre le nombre de défaillances observées et attendues dans les intervalles  $I_1, \dots, I_k$ . La méthode d'estimation du maximum de vraisemblance, dont les propriétés asymptotiques sont bien connues, a été utilisée pour déterminer les différents paramètres des modèles.

Les auteurs ont procédé comme suit :

Soit  $X_i = (X_1, \dots, X_n)$  un échantillon de variables aléatoires censurées à droite, tel que :

$$(X_1, \delta_1), \dots, (X_n, \delta_n), \quad (1)$$

Où :

$$X_i = T_i \wedge C_i, \quad \delta_i = 1_{\{T_i \leq C_i\}}.$$

En supposant que les taux de panne  $T_1, \dots, T_n$  sont absolument continus indépendants et identiquement distribués, *i.i.d.* et en notant l'échantillon (1) en termes de processus stochastiques  $N_i(t)$  et  $Y_i(t)$ . On a :

$$(N_1(t), Y_1(t), t \geq 0, \dots, N_n(t), Y_n(t), t \geq 0).$$

Où :

Les trajectoires du processus de comptage  $N_i$  ont la forme :

$$N_i(t) = \begin{cases} 0, & 0 \leq t < X_i \\ 1, & t \geq X_i \end{cases}.$$

Les trajectoires du processus stochastique  $Y_i$  ont la forme :

$$Y_i(t) = \begin{cases} 0, & t > X_i \\ 1, & 0 \leq t \leq X_i \end{cases}.$$

En posant :

$$N(t) = \sum_{i=1}^n N_i(t) \quad \text{et} \quad Y(t) = \sum_{i=1}^n Y_i(t),$$

en utilisant ces processus, ils obtiennent les relations suivantes :

$$\int_0^\infty \log h(u, \theta) dN_i(u) = \begin{cases} \ln \lambda(X_i, \theta), & \delta_i = 1 \\ 0, & \delta_i = 0 \end{cases} = \delta_i h(X_i, \theta),$$

et

$$\int_0^\infty Y_i(u) \log h(u, \theta) d(u) = \int_0^{X_i} h(u, \theta) d(u) = -\ln S(X_i, \theta); \quad \{\theta \in \Theta\}.$$

En considérant l'hypothèse nulle  $H_0$

$$H_0 : F(x) \in \mathcal{F}_0 = \{F_0(x, \theta); \theta \in \Theta \subset \mathbb{R}^s\},$$

où  $\theta = (\theta_1, \dots, \theta_s)^T \in \Theta \subset \mathbb{R}^s$  est le vecteur de paramètres inconnus et  $F_0$  est une fonction de répartition connue. En groupant les données initiales observées notées en termes des processus de comptages  $N_i(t)$  et  $Y_i(t)$  dans un intervalle de temps  $[0, \tau]$  en  $k > s$  sous-intervalles :

$$I_j = (a_{j-1}, a_j], \quad j = \overline{1, k}, \quad a_0 = 0, \quad a_k = \max(X_{(l)}, \tau),$$

où  $X_{(i)}$  est le  $i$ ème élément de la statistique d'ordre  $X_{(1)}, \dots, (i_n)$ . Et en notant par :

$$U_j = N(a_j) - N(a_{j-1}) = \sum_{i: X_i \in I_j} \delta_i, \quad j = \overline{1, k}$$

le nombre de défaillances observées dans le  $j^{ième}$  intervalle, sous l'hypothèse  $H_0$ , on estime le nombre de défaillances théorique  $e_j$  dans l'intervalle  $I_j$ ; en tenant compte de l'équation :

$$E[N(t)] = E \int_0^t h(u, \hat{\theta}) Y(u) du,$$

$\hat{\theta}$  est l'estimateur du maximum de vraisemblance du paramètre  $\theta$ . Alors, le nombre de défaillances théoriques peut être donné par :

$$e_j = E \int_{a_{j-1}}^{a_j} h(u, \hat{\theta}) Y(u) du = \sum_{i: t_i > a_{j-1}} [H(a_j \wedge t_i, \hat{\theta}) - H(a_{j-1}, \hat{\theta})],$$

où  $a \wedge b = \min(a, b)$ . La statistique du test de chi-deux de *Bagdonavicius et Nikulin* (2011) est basée sur la statistique suivante :

$$Y_n^2 = Z^T \hat{V}^{-1} Z.$$

Où le vecteur  $Z$  est donné par :

$$Z = (Z_1, \dots, Z_k)^T; \quad Z_j = \frac{1}{\sqrt{n}} (U_j - e_j), \quad j = \overline{1, k}.$$

$\hat{V}^-$  est une matrice inverse généralisée de  $\hat{V}$  donnée par :

$$\hat{V}^- = \hat{A}^{-1} + \hat{A}^{-1} \hat{C}^T \hat{G}^- \hat{C} \hat{A}^{-1}, \quad \hat{G}^- = \hat{i} - \hat{C} \hat{A}^{-1} \hat{C}^T,$$

où :

$$\hat{A} = \begin{pmatrix} \hat{A}_1 & 0 & \dots & 0 \\ 0 & \hat{A}_2 & \dots & 0 \\ 0 & 0 & \dots & 0 \\ 0 & 0 & \dots & \hat{A}_k \end{pmatrix} ; \quad \hat{C} = \begin{pmatrix} \hat{C}_{11} & \hat{C}_{12} & \dots & \hat{C}_{1k} \\ \hat{C}_{21} & \hat{C}_{22} & \dots & \hat{C}_k \\ \vdots & \vdots & & \vdots \\ \hat{C}_{s1} & \hat{C}_{s2} & & \hat{C}_{sk} \end{pmatrix}.$$

Alors,

$$\hat{A}_j = \frac{U_j}{n} \quad \text{et} \quad \hat{C}_j = \frac{1}{n} \sum_{i: X_i \in I_j} \frac{\partial}{\partial \theta} \ln h(X_i, \theta).$$

D'où la statistique du test  $Y_n^2$  peut être écrite sous la forme :

$$Y_n^2 = Z^T \hat{A}^{-1} Z + Z^T \hat{A}^{-1} \hat{C}^T \hat{G}^- \hat{C} \hat{A}^{-1} Z,$$

$$Y_n^2 = \sum_{i=1}^n \frac{(U_j - e_j)^2}{U_j} + Q,$$

où,

$$Q = W^T \hat{G}^- W, \quad W = \hat{C} \hat{A}^{-1} Z = (W_1, \dots, W_s)^T,$$

$$\hat{G} = [\hat{g}_{ll'}]_{s \times s}, \quad \hat{g}_{ll'} = \hat{i}_{ll'} - \sum_{j=1}^k \hat{C}_{lj} \hat{C}_{l'j} \hat{A}_j^{-1}, \quad W_l = \sum_{j=1}^k \hat{C}_{lj} \hat{A}_j^{-1} Z_j,$$

$$\hat{i}_{ll'} = \frac{1}{n} \sum_{j=1}^n \delta_i \frac{\partial \ln \lambda(X_i, \hat{\theta})}{\partial \theta_l} \frac{\partial \ln \lambda(X_i, \hat{\theta})}{\partial \theta_{l'}}, \quad \hat{\theta} = (\theta_1, \dots, \theta_s),$$

$$Z_j = \frac{1}{\sqrt{n}} (U_j - e_j), U_j = \sum_{i: X_i \in I_j} \delta_i, \quad j = \overline{1, k}, \quad l, l' = \overline{1, s}.$$

Sous l'hypothèse  $H_0$ , la distribution limite de la statistique de test  $Y_n^2$  est une distribution chi-deux à  $\nu$  degrés de liberté avec

$$\nu = \text{rank}(V^-) = \text{Tr}(V^- V).$$

Si  $G$  est une matrice non-dégénérée alors  $\nu = k$ .

L'hypothèse nulle  $H_0$  est rejetée pour un risque d'erreur  $\epsilon$  fixé, si  $Y_n^2 > \chi_c^2(\nu)$ .

### Choix des intervalles de regroupement des données :

Afin d'assurer l'égalité des effectifs théoriques des taux de pannes, les auteurs ont considéré les limites  $a_j$  des intervalles de regroupement comme étant des fonctions aléatoires. L'idée consiste à diviser l'intervalle  $[0, \tau]$  en  $k$  intervalles, tels que :

$$b_i = (n - i) H(X_{(i)}, \hat{\theta}) + \sum_{l=1}^i H(X_{(l)}, \hat{\theta}),$$

où  $X_{(i)}$  est le  $i$ ème élément de la statistique d'ordre  $(X_{(1)}, \dots, X_{(n)})$ .

Si  $i$  le plus petit nombre naturel vérifiant :

$$E_j \in [b_{i-1}, b_i], \quad j = \overline{1, k},$$

alors,

$$E_j = (n - i + 1) H(a, \hat{\theta}) + \sum_{l=1}^{i-1} H(X_{(l)}, \hat{\theta}).$$

D'où les limites  $\hat{a}_j$  sont obtenues par :

$$\hat{a}_j = H^{-1} \left\{ \frac{E_j - \sum_{l=1}^{i-1} H(X_{(l)}, \hat{\theta})}{n - i + 1}, \hat{\theta} \right\}, \hat{a}_k = \max(X_{(n)}, \tau).$$

Où  $H^{-1}$  est l'inverse de la fonction du taux de hasard cumulé  $H$ . On a  $0 < \hat{a}_1 < \hat{a}_2 < \dots < \hat{a}_k = \tau$ .

Pour ce choix d'intervalles, on a :

$$e_j = \frac{E_k}{k}, \forall j,$$

tel que  $E_k = \sum_{i=1}^n H(X_i, \hat{\theta})$ .

## Le modèle N-Weibull-Weibull (*NWW*)

En théorie des probabilités, la loi de Weibull, nommée d'après le mathématicien *Waldoddi Weibull* (1951), est une loi de probabilité continue. c'est la distribution la plus utilisée pour modéliser les durées de vie et le temps de défaillance dans les tests de vie, de survie et la théorie de fiabilité. Son importance est liée à sa flexibilité et ses divers domaines d'application, nous citons l'analyse de survie, l'étude de fiabilité, la théorie des valeurs extrêmes, l'économie et elle peut modéliser aussi la taille des sinistres en réassurance et en assurance générale. Elle est aussi utile en génie électrique, ingénierie industrielle, dans les prévisions météorologiques, la dispersion des signaux dans les systèmes radar, l'industrie éolienne pour décrire les distributions de vitesse du vent et en ingénierie des systèmes de communication.

En général, en théorie de fiabilité et en analyse de survie, les données observées ont souvent un comportement non-monotone, ce qui n'est pas le cas de la distribution Weibull classique, c'est pourquoi de nouvelles généralisations ont été proposées. Parmi celles-ci, on s'est intéressé à une distribution à quatre (4) paramètres, appelée la distribution New-Weibull-Weibull (*NWW*). L'avantage de la distribution (*NWW*) est la flexibilité de sa fonction du taux de défaillance qui peut prendre diverses formes telles que

la forme baignoire, baignoire inversée (unimodale) et autres selon la valeur des paramètres.

Dans ce chapitre, on commence en premier lieu par présenter cette nouvelle distribution *NWW* et établir ses caractéristiques ainsi qu'une représentation graphique des différentes formes du taux de hasard. Par la suite, on construit des statistiques de test d'ajustement pour valider le choix de ce modèle. L'étude a été menée aussi bien pour le cas de données complètes que pour les données censurées à droite. Des simulations numériques et des applications à des données réelles sont utilisées pour confirmer les résultats obtenus.

## 1 Présentation du modèle *NWW*

En utilisant la généralisation introduite par *Tahir et al.*, (2016) et présentée dans le chapitre 1, section 2, et en remplaçant la distribution de base  $G(x)$  par la fonction de répartition de la distribution Weibull, on peut caractériser la distribution New-Weibull-Weibull (*NWW*) par la fonction de répartition (cdf) et la densité de probabilité (pdf), données, respectivement, par :

$$F_{NWW}(x, \vartheta) = \exp \left\{ -\alpha \left[ -\log(1 - \exp(-\lambda x^\gamma)) \right]^\xi \right\}, x > 0, \vartheta = (\alpha, \xi, \lambda, \gamma) > 0,$$

$$\begin{aligned} f_{NWW}(x, \vartheta) &= \alpha \xi \frac{\lambda \gamma x^{\gamma-1} \exp(-\lambda x^\gamma)}{1 - \exp(-\lambda x^\gamma)} \left\{ -\log[1 - \exp(-\lambda x^\gamma)] \right\}^{\xi-1} \\ &\quad \times \exp \left( -\alpha \left\{ -\log[1 - \exp(-\lambda x^\gamma)] \right\}^\xi \right). \end{aligned}$$

Les fonctions de survie  $S(x)$ , du taux de hasard (hrf)  $h(x)$  et du taux de hasard cumulé  $H(x)$  s'écrivent respectivement :

$$S_{NWW}(x, \vartheta) = 1 - \exp \left\{ -\alpha \left[ -\log(1 - \exp(-\lambda x^\gamma)) \right]^\xi \right\}.$$

$$h_{NWW}(x, \vartheta) = \frac{\alpha \xi \frac{\lambda \gamma x^{\gamma-1} \exp(-\lambda x^\gamma)}{1 - \exp(-\lambda x^\gamma)} \left\{ -\log[1 - \exp(-\lambda x^\gamma)] \right\}^{\xi-1}}{1 - \exp \left( -\alpha \left\{ -\log[1 - \exp(-\lambda x^\gamma)] \right\}^\xi \right)} \times \exp \left( -\alpha \left\{ -\log[1 - \exp(-\lambda x^\gamma)] \right\}^\xi \right),$$

$$H_{NWW}(x, \vartheta) = -\log \left\{ 1 - \exp \left( -\alpha \left[ -\log(1 - \exp(-\lambda x^\gamma)) \right]^\xi \right) \right\}.$$

Selon les valeurs prises par les paramètres, on présente les différentes formes des fonctions de densité de probabilité et du taux de hasard de cette nouvelle distribution  $NWW$ , en comparaison respectivement avec celles de la distribution Weibull-Weibull ( $WW$ ), voir la Figure 2.1 et la Figure 2.2

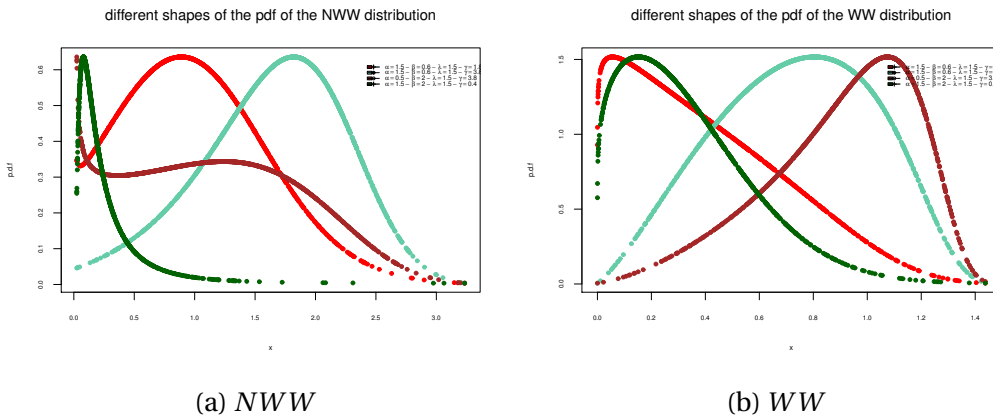


FIGURE 2.1 – pdf du modèle  $NWW$  vs.  $WW$ .

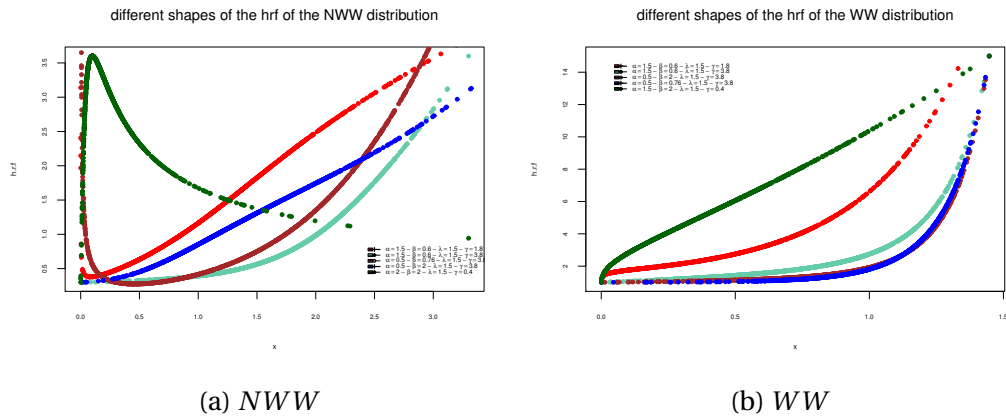


FIGURE 2.2 – hrf du modèle *NWW* vs. *WW*.

Selon la Figure 2.1, la distribution *NWW* produit une variété de formes de densité de probabilité, y compris des formes asymétriques à gauche, à droite, en baignoire et en cloche.

La Figure 2.2 montre également la capacité de la famille à générer une variété de formes du taux de hasard, y compris croissante, baignoire, cloche et la forme en J. D'où, la distribution *NWW* peut être très utile pour ajuster divers ensembles de données avec des formes variées, telles que des données économiques et de fiabilité.

## 2 Estimation

### 2.1 Estimation des paramètres inconnus dans le cas complet

Vu ses propriétés de convergence et de normalité asymptotique, on a utilisé la méthode d'estimation du maximum de vraisemblance pour estimer les paramètres inconnus du modèle *NWW*.

Considérons  $X = (X_1, \dots, X_n)^T$  un échantillon de variables aléatoires complètes indépendantes et identiquement distribuées (*i.i.d.*) de la distribution *NWW*, de vecteur de

paramètres  $\vartheta = (\alpha, \xi, \lambda, \gamma)$ , la fonction de vraisemblance de  $X$  s'écrit :

$$\begin{aligned} L_{NWW}(x_i, \vartheta) &= \prod_{i=1}^n f_{NWW}(x_i, \vartheta) \\ &= \prod_{i=1}^n \left[ \alpha \xi \frac{\lambda \gamma x_i^{\gamma-1} \exp(-\lambda x_i^\gamma)}{1 - \exp(-\lambda x_i^\gamma)} \{-\log[1 - \exp(-\lambda x_i^\gamma)]\}^{\xi-1} \exp\left(-\alpha \{-\log[1 - \exp(-\lambda x_i^\gamma)]\}^\xi\right) \right]. \end{aligned}$$

L'expression de la fonction de log-vraisemblance s'écrit alors :

$$\begin{aligned} l(x_i, \vartheta) &= \sum_{i=1}^n \log f_{NWW}(x_i, \alpha, \xi, \lambda, \gamma) \\ &= n \log(\alpha \xi \lambda \gamma) + \sum_{i=1}^n \log(x_i^{\gamma-1}) - \sum_{i=1}^n (\lambda x_i^\gamma) + (\xi - 1) \sum_{i=1}^n \log\{-\log[1 - \exp(-\lambda x_i^\gamma)]\} \\ &\quad - \alpha \sum_{i=1}^n \{-\log[1 - \exp(-\lambda x_i^\gamma)]\}^\xi - \sum_{i=1}^n \log[1 - \exp(-\lambda x_i^\gamma)]. \end{aligned}$$

Les fonctions scores sont obtenues en dérivant la fonction log-vraisemblance par rapport au vecteur des paramètres  $\vartheta$  et sont données par :

$$\begin{aligned} \frac{\partial l_{NWW}(x_i, \vartheta)}{\partial \alpha} &= \frac{n}{\alpha} - \sum_{i=1}^n [-\log(1 - \exp(-\lambda x_i^\gamma))]^\xi, \\ \frac{\partial l_{NWW}(x_i, \vartheta)}{\partial \xi} &= \frac{n}{\xi} + \sum_{i=1}^n \log[-\log(1 - \exp(-\lambda x_i^\gamma))] - \alpha \sum_{i=1}^n \log[K_{NWW}(x_i, \vartheta)] \times K_{NWW}^\xi(x_i, \vartheta), \\ \frac{\partial l_{NWW}(x_i, \vartheta)}{\partial \lambda} &= \frac{n}{\lambda} - \sum_{i=1}^n x_i^\gamma + (\xi - 1) \sum_{i=1}^n \frac{x_i^\gamma \exp(-\lambda x_i^\gamma)}{Z_{NWW}(x_i, \vartheta)} \\ &\quad + \alpha \xi \sum_{i=1}^n \frac{x_i^\gamma \exp(-\lambda x_i^\gamma) K_{NWW}^{\xi-1}(x_i, \vartheta)}{(1 - \exp(-\lambda x_i^\gamma))} - \sum_{i=1}^n \frac{x_i^\gamma \exp(-\lambda x_i^\gamma)}{(1 - \exp(-\lambda x_i^\gamma))}, \\ \frac{\partial l_{NWW}(x_i, \vartheta)}{\partial \gamma} &= \frac{n}{\gamma} + \sum_{i=1}^n \log(x_i) - \lambda \sum_{i=1}^n x_i^\gamma \log(x_i) + \lambda (\xi - 1) \sum_{i=1}^n \frac{M_{NWW}(x_i, \vartheta)}{Z_{NWW}(x_i, \vartheta)} \\ &\quad - \lambda \sum_{i=1}^n \frac{M_{NWW}(x_i, \vartheta)}{(1 - \exp(-\lambda x_i^\gamma))} \left[ 1 - \alpha \xi K_{NWW}^{\xi-1}(x_i, \vartheta) \right], \end{aligned}$$

où :

$$\begin{aligned} K_{NWW}(x, \vartheta) &= \{-\log[1 - \exp(-\lambda x^\gamma)]\}, \\ M_{NWW}(x, \vartheta) &= x^\gamma \times \log(x) \times \exp(-\lambda x^\gamma), \\ Z_{NWW}(x, \vartheta) &= [1 - \exp(-\lambda x^\gamma)] \times \log[1 - \exp(-\lambda x^\gamma)]. \end{aligned}$$

Les estimateurs du maximum de vraisemblance (EMV) sont obtenus en égalant à zéro les dérivées premières ci-dessus. Pour résoudre ce système d'équations non linéaires, des méthodes numériques sont nécessaires telles que la méthode de *Newton-Raphson*, la méthode de Monte Carlo, l'algorithme BB ou autres.

## 2.2 La matrice d'information de Fisher estimée dans le cas complet

Les éléments de la matrice d'information de Fisher  $\hat{I} = [i_{ll'}]_{s \times s}$  obtenus à partir des dérivées secondes de la fonction de log-vraisemblance par rapport aux paramètres, sont donnés par :

$$\begin{aligned} \hat{i}_{12} &= -\sum_{i=1}^n \log[K_{NWW}(x_i, \hat{\vartheta})] \times [K_{NWW}(x_i, \hat{\vartheta})]^{\hat{\xi}}, \\ \hat{i}_{13} &= \hat{\xi} \sum_{i=1}^n \frac{x_i^{\hat{\gamma}} \exp(-\hat{\lambda} x_i^{\hat{\gamma}}) [K_{NWW}(x_i, \hat{\vartheta})]^{\hat{\xi}-1}}{(1 - \exp(-\hat{\lambda} x_i^{\hat{\gamma}}))}, \\ \hat{i}_{14} &= \hat{\xi} \hat{\lambda} \sum_{i=1}^n \frac{M_{NWW}(x_i, \hat{\vartheta}) \times [K_{NWW}(x_i, \hat{\vartheta})]^{\hat{\xi}-1}}{(1 - \exp(-\hat{\lambda} x_i^{\hat{\gamma}}))}, \\ \hat{i}_{23} &= \hat{\alpha} \hat{\xi} \sum_{i=1}^n \frac{x_i^{\hat{\gamma}} \exp(-\hat{\lambda} x_i^{\hat{\gamma}}) \log[K_{NWW}(x_i, \hat{\vartheta})] \times [K_{NWW}(x_i, \hat{\vartheta})]^{\hat{\xi}-1}}{1 - \exp(-\hat{\lambda} x_i^{\hat{\gamma}})}, \\ &\quad + \sum_{i=1}^n x_i^{\hat{\gamma}} \exp(-\hat{\lambda} x_i^{\hat{\gamma}}) \left\{ \frac{1 - \hat{\alpha} [K_{NWW}(x_i, \hat{\vartheta})]^{\hat{\xi}}}{Z_{NWW}(x_i, \hat{\vartheta})} \right\}, \end{aligned}$$

$$\begin{aligned}
 \hat{i}_{24} &= \hat{\lambda} \sum_{i=1}^n \frac{M_{NWW}(x_i, \hat{\vartheta})}{Z_{NWW}(x_i, \hat{\vartheta})} \left[ 1 - \hat{\alpha} [K_{NWW}(x_i, \hat{\vartheta})]^{\hat{\xi}} \right] \\
 &\quad + \hat{\alpha} \hat{\xi} \hat{\lambda} \sum_{i=1}^n \frac{M_{NWW}(x_i, \hat{\vartheta}) \log [K_{NWW}(x_i, \hat{\vartheta})] [K_{NWW}(x_i, \hat{\vartheta})]^{\hat{\xi}-1}}{1 - \exp(-\hat{\lambda} x_i^{\hat{\gamma}})}, \\
 \hat{i}_{34} &= \hat{\lambda} \sum_{i=1}^n \frac{x_i^{2\hat{\gamma}} \log(x_i) \exp(-2\hat{\lambda} x_i^{\hat{\gamma}})}{[1 - \exp(-\hat{\lambda} x_i^{\hat{\gamma}})]^2} \left\{ 1 - \hat{\alpha} \hat{\xi} [K_{NWW}^{\hat{\xi}-1}(x_i, \hat{\vartheta}) + (\hat{\xi}-1) K_{NWW}^{\hat{\xi}-2}(x_i, \hat{\vartheta})] \right\} \\
 &\quad - \sum_{i=1}^n x_i^{\hat{\gamma}} \log(x_i) + \sum_{i=1}^n [1 - \hat{\lambda} x_i^{\hat{\gamma}}] \frac{M_{NWW}(x_i, \hat{\vartheta})}{1 - \exp(-\hat{\lambda} x_i^{\hat{\gamma}})} \left[ \hat{\alpha} \hat{\xi} K_{NWW}^{\hat{\xi}-1}(x_i, \hat{\vartheta}) - 1 \right] \\
 &\quad + (\hat{\xi}-1) \sum_{i=1}^n \frac{M_{NWW}(x_i, \hat{\vartheta})}{Z_{NWW}(x_i, \hat{\vartheta})} \left[ 1 - \hat{\lambda} x_i^{\hat{\gamma}} - \frac{\hat{\lambda} x_i^{\hat{\gamma}} \exp(-\hat{\lambda} x_i^{\hat{\gamma}})}{1 - \exp(-\hat{\lambda} x_i^{\hat{\gamma}})} - \frac{\hat{\lambda} x_i^{\hat{\gamma}} \exp(-\hat{\lambda} x_i^{\hat{\gamma}})}{Z_{NWW}(x_i, \hat{\vartheta})} \right], \\
 \hat{i}_{11} &= -\frac{n}{\hat{\alpha}^2}, \\
 \hat{i}_{22} &= -\frac{n}{\hat{\xi}^2} - \hat{\alpha} \sum_{i=1}^n \log^2 [K_{NWW}(x_i, \hat{\vartheta})] \times [K_{NWW}(x_i, \hat{\vartheta})]^{\hat{\xi}}, \\
 \hat{i}_{33} &= -\frac{n}{\hat{\lambda}^2} + \sum_{i=1}^n \frac{x_i^{2\hat{\gamma}} \exp(-\hat{\lambda} x_i^{\hat{\gamma}})}{1 - \exp(-\hat{\lambda} x_i^{\hat{\gamma}})} \left[ 1 - \hat{\alpha} \hat{\xi} [K_{NWW}(x_i, \hat{\vartheta})]^{\hat{\xi}-1} \right] \\
 &\quad + \sum_{i=1}^n \frac{x_i^{2\hat{\gamma}} \exp(-2\hat{\lambda} x_i^{\hat{\gamma}})}{[1 - \exp(-\hat{\lambda} x_i^{\hat{\gamma}})]^2} \left\{ 1 - \hat{\alpha} \hat{\xi} K_{NWW}^{\hat{\xi}-1}(x_i, \hat{\vartheta}) \left[ 1 + \frac{(\hat{\xi}-1)}{K_{NWW}(x_i, \hat{\vartheta})} \right] \right\} \\
 &\quad - (\hat{\xi}-1) \sum_{i=1}^n \frac{x_i^{2\hat{\gamma}} \exp(-\hat{\lambda} x_i^{\hat{\gamma}})}{Z_{NWW}(x_i, \hat{\vartheta})} \left[ 1 + \frac{\exp(-\hat{\lambda} x_i^{\hat{\gamma}})}{1 - \exp(-\hat{\lambda} x_i^{\hat{\gamma}})} + \frac{\exp(-\hat{\lambda} x_i^{\hat{\gamma}})}{Z_{NWW}(x_i, \hat{\vartheta})} \right], \\
 \hat{i}_{44} &= -\frac{n}{\hat{\gamma}^2} - \hat{\lambda} \sum_{i=1}^n x_i^{\hat{\gamma}} \log^2(x_i) - \hat{\lambda} (\hat{\xi}-1) \sum_{i=1}^n \frac{M_{NWW}(x_i, \hat{\vartheta}) \log(x_i)}{Z_{NWW}(x_i, \hat{\vartheta})} [\hat{\lambda} x_i^{\hat{\gamma}} - 1] \\
 &\quad + \hat{\lambda} \sum_{i=1}^n \frac{1 - \hat{\alpha} \hat{\xi} K_{NWW}^{\hat{\xi}-1}(x_i, \hat{\vartheta})}{1 - \exp(-\hat{\lambda} x_i^{\hat{\gamma}})} \left[ \hat{\lambda} x_i^{2\hat{\gamma}} \log^2(x_i) \exp(-\hat{\lambda} x_i^{\hat{\gamma}}) - M_{NWW}(x_i, \hat{\vartheta}) \log(x_i) \right] \\
 &\quad - \hat{\lambda}^2 (\hat{\xi}-1) \sum_{i=1}^n \frac{M_{NWW}^2(x_i, \hat{\vartheta})}{Z_{NWW}(x_i, \hat{\vartheta})} \left[ \frac{1}{1 - \exp(-\hat{\lambda} x_i^{\hat{\gamma}})} + \frac{1}{Z_{NWW}(x_i, \hat{\vartheta})} \right] \\
 &\quad + \hat{\lambda}^2 \sum_{i=1}^n \frac{M_{NWW}^2(x_i, \hat{\vartheta})}{[1 - \exp(-\hat{\lambda} x_i^{\hat{\gamma}})]^2} \left[ 1 - \hat{\alpha} \hat{\xi} K_{NWW}^{\hat{\xi}-1}(x_i, \hat{\vartheta}) - \hat{\alpha} (\hat{\xi}-1) \hat{\xi} K_{NWW}^{\hat{\xi}-2}(x_i, \hat{\vartheta}) \right].
 \end{aligned}$$

### 2.3 Estimation des paramètres inconnus dans le cas censuré

Soit  $X = (X_1, \dots, X_n)$  un  $n$ -échantillon aléatoire censuré à droite et distribué selon la distribution N-Weibull-Weibull de vecteur de paramètres inconnus  $\vartheta = (\alpha, \xi, \lambda, \gamma)^T$  :

$$X_i = T_i \wedge C_i, \quad \delta_i = 1_{\{T_i \leq C_i\}},$$

où  $T_i$  représente les taux de défaillances qui sont des variables aléatoires absolument continues et *i.i.d.* et  $C_i$  désignent les temps de censure à droite, tels que  $T_i$  et  $C_i$  sont des variables indépendantes.

On détermine la fonction de vraisemblance de la distribution N-Weibull-Weibull par l'équation suivante :

$$\begin{aligned} L_{NWW}(x, \vartheta) &= \prod_{i=1}^n h_{NWW}^{\delta_i}(x_i, \vartheta) S_{NWW}(x_i, \vartheta) \\ &= \prod_{i=1}^n \left[ \frac{\alpha \xi \frac{\lambda \gamma x_i^{\gamma-1} \exp(-\lambda x_i^\gamma)}{1 - \exp(-\lambda x_i^\gamma)} \{-\log[1 - \exp(-\lambda x_i^\gamma)]\}^{\xi-1} \times \exp\left(-\alpha \{-\log[1 - \exp(-\lambda x_i^\gamma)]\}^\xi\right)}{1 - \exp\left(-\alpha \{-\log[1 - \exp(-\lambda x_i^\gamma)]\}^\xi\right)} \right]^{\delta_i} \\ &\quad \times \left[ 1 - \exp\left(-\alpha \{-\log[1 - \exp(-\lambda x_i^\gamma)]\}^\xi\right) \right], \end{aligned}$$

$\delta_i = 1_{\{T_i \leq C_i\}}$  est l'indicateur de censure. Alors, la fonction log-vraisemblance de la distribution NWW s'écrit :

$$\begin{aligned} l_{NWW}(x, \vartheta) &= \sum_{i=1}^n \delta_i \log h(x_i, \vartheta) + \sum_{i=1}^n \log S(x_i, \vartheta) \\ &= r \log(\alpha \xi \lambda \gamma) + (\gamma - 1) \sum_{i \in F} \log(x_i) - \lambda \sum_{i \in F} x_i^\gamma + \sum_{i \in C} \log \left[ 1 - \exp\left(-\alpha \{-\log[1 - \exp(-\lambda x_i^\gamma)]\}^\xi\right) \right] \\ &\quad - \alpha \sum_{i \in F} \{-\log[1 - \exp(-\lambda x_i^\gamma)]\}^\xi - \sum_{i \in F} \log[1 - \exp(-\lambda x_i^\gamma)] + (\xi - 1) \sum_{i \in F} \log \{-\log[1 - \exp(-\lambda x_i^\gamma)]\} \end{aligned}$$

Où  $r$  est le nombre de défaillances et  $F$  et  $C$  désignent respectivement les ensembles d'observations non censurées et censurées à droite. Les fonctions scores du vecteur des paramètres inconnus du modèle N-Weibull-Weibull sont données par les formules suivantes :

$$\begin{aligned} \frac{\partial l_{NWW}(x_i, \vartheta)}{\partial \alpha} &= \frac{r}{\alpha} - \sum_{i \in F} K_{NWW}^{\xi}(x_i, \vartheta) + \sum_{i \in C} \frac{K_{NWW}^{\xi}(x_i, \vartheta) \exp[-\alpha K_{NWW}^{\xi}(x_i, \vartheta)]}{1 - \exp[-\alpha K_{NWW}^{\xi}(x_i, \vartheta)]}, \\ \frac{\partial l_{NWW}(x_i, \vartheta)}{\partial \xi} &= \frac{r}{\xi} + \sum_{i \in F} \log[K_{NWW}(x_i, \vartheta)] - \alpha \sum_{i \in F} \log[K_{NWW}(x_i, \vartheta)] K_{NWW}^{\xi}(x_i, \vartheta) \\ &\quad + \alpha \sum_{i \in C} \frac{\log[K_{NWW}(x_i, \vartheta)] K_{NWW}^{\xi}(x_i, \vartheta) \exp[-\alpha K_{NWW}^{\xi}(x_i, \vartheta)]}{1 - \exp[-\alpha K_{NWW}^{\xi}(x_i, \vartheta)]}, \\ \frac{\partial l_{NWW}(x_i, \vartheta)}{\partial \lambda} &= \frac{r}{\lambda} - \sum_{i \in F} x_i^{\gamma} - \sum_{i \in F} \frac{x_i^{\gamma} \exp(-\lambda x_i^{\gamma})}{1 - \exp(-\lambda x_i^{\gamma})} + (\xi - 1) \sum_{i \in F} \frac{x_i^{\gamma} \exp(-\lambda x_i^{\gamma})}{Z_{NWW}(x_i, \vartheta)} \\ &\quad + \alpha \xi \frac{x_i^{\gamma} K_{NWW}^{\xi-1}(x_i, \vartheta)}{1 - \exp(-\lambda x_i^{\gamma})} \left[ \sum_{i \in F} \exp(-\lambda x_i^{\gamma}) - \sum_{i \in C} \frac{M_{NWW}(x_i, \vartheta)}{1 - \exp[-\alpha K_{NWW}^{\xi}(x_i, \vartheta)]} \right], \\ \frac{\partial l_{NWW}(x_i, \vartheta)}{\partial \gamma} &= \frac{r}{\gamma} + \sum_{i \in F} \log(x_i) [1 - \lambda x_i^{\gamma}] + \lambda (\xi - 1) \sum_{i \in F} \frac{M_{NWW}(x_i, \vartheta)}{Z_{NWW}(x_i, \vartheta)} \\ &\quad + \lambda \sum_{i \in F} \frac{M_{NWW}(x_i, \vartheta)}{1 - \exp(-\lambda x_i^{\gamma})} [\alpha \xi K_{NWW}^{\xi-1}(x_i, \vartheta) - 1] - \alpha \xi \lambda \sum_{i \in C} \frac{x_i^{\gamma} \log(x_i) K_{NWW}^{\xi-1}(x_i, \vartheta) D_{NWW}(x_i, \vartheta)}{1 - \exp(-\lambda x_i^{\gamma}) 1 - \exp[-\alpha K_{NWW}^{\xi}(x_i, \vartheta)]}. \end{aligned}$$

Où :

$$\begin{aligned} D_{NWW}(x, \vartheta) &= \exp\left(-\alpha \{-\log[1 - \exp(-\lambda x^{\gamma})]\}^{\xi} - \lambda x^{\gamma}\right) \\ K_{NWW}(x, \vartheta) &= \{-\log[1 - \exp(-\lambda x^{\gamma})]\}, \\ M_{NWW}(x, \vartheta) &= x^{\gamma} \times \log(x) \times \exp(-\lambda x^{\gamma}), \\ Z_{NWW}(x, \vartheta) &= [1 - \exp(-\lambda x^{\gamma})] \times \log[1 - \exp(-\lambda x^{\gamma})]. \end{aligned}$$

De même que pour le cas des données complètes, on peut utiliser différentes méthodes numériques pour résoudre ce système d'équations.

## 2.4 La matrice d'information de Fisher estimée dans le cas censuré

On a :

$$i_{ll'} = \frac{1}{n} \sum_{i=1}^n \frac{\partial \log h(x_i, \hat{\vartheta})}{\partial \hat{\vartheta}_l} \frac{\partial \log h(x_i, \hat{\vartheta})}{\partial \hat{\vartheta}_{l'}}.$$

Alors, les éléments de la matrice d'information de Fisher  $\hat{T}$  sont :

$$\begin{aligned} \hat{i}_{11} &= \frac{1}{n} \sum_{i=1}^n \delta i \left[ \frac{-1}{\hat{\alpha}^2} + \frac{K_{NWW}^{2\hat{\xi}}(x_i, \hat{\vartheta}) \exp[-\hat{\alpha} K_{NWW}^{\hat{\xi}}(x_i, \hat{\vartheta})]}{\{1 - \exp[-\hat{\alpha} K_{NWW}^{\hat{\xi}}(x_i, \hat{\vartheta})]\}^2} \right], \\ \hat{i}_{22} &= \frac{1}{n} \sum_{i=1}^n \delta i \left[ \frac{-1}{\hat{\xi}^2} - \hat{\alpha} \frac{\log^2 [K_{NWW}(x_i, \hat{\vartheta})] K_{NWW}^{\hat{\xi}}(x_i, \hat{\vartheta})}{1 - \exp[-\hat{\alpha} K_{NWW}^{\hat{\xi}}(x_i, \hat{\vartheta})]} \right] + \\ &\quad \frac{1}{n} \sum_{i=1}^n \delta i \left[ \hat{\alpha}^2 \frac{\log^2 [K_{NWW}(x_i, \hat{\vartheta})] K_{NWW}^{2\hat{\xi}}(x_i, \hat{\vartheta}) \exp[-\hat{\alpha} K_{NWW}^{\hat{\xi}}(x_i, \hat{\vartheta})]}{\{1 - \exp[-\hat{\alpha} K_{NWW}^{\hat{\xi}}(x_i, \hat{\vartheta})]\}^2} \right], \\ \hat{i}_{33} &= \frac{1}{n} \sum_{i=1}^n \delta i \left\{ \frac{-1}{\hat{\lambda}^2} + \frac{x_i^{2\hat{\gamma}} \exp(-\hat{\lambda} x_i^{\hat{\gamma}})}{1 - \exp(-\hat{\lambda} x_i^{\hat{\gamma}})} \left[ 1 - \frac{(\hat{\xi} - 1)}{\log[1 - \exp(-\hat{\lambda} x_i^{\hat{\gamma}})]} - \hat{\alpha} \hat{\xi} \frac{K_{NWW}^{\hat{\xi}-1}(x_i, \hat{\vartheta})}{1 - \exp(-\hat{\lambda} x_i^{\hat{\gamma}})} \right] \right\} \\ &\quad - \frac{(\hat{\xi} - 1)}{n} \sum_{i=1}^n \delta i \frac{x_i^{2\hat{\gamma}} \exp(-2\hat{\lambda} x_i^{\hat{\gamma}})}{D_{NWW}(x_i, \hat{\vartheta})} \left[ \frac{1}{D_{NWW}(x_i, \hat{\vartheta})} + \frac{1}{1 - \exp(-\hat{\lambda} x_i^{\hat{\gamma}})} \right] + \\ &\quad \frac{\hat{\alpha}^2 \hat{\xi}^2}{n} \sum_{i=1}^n \delta i \frac{K_{NWW}^{2\hat{\xi}-2}(x_i, \hat{\vartheta}) D_{NWW}(x_i, \hat{\vartheta})}{V_{NWW}(x_i, \hat{\vartheta})} \left[ \frac{1}{1 - \exp(-\hat{\lambda} x_i^{\hat{\gamma}})} + \frac{M_{NWW}(x_i, \hat{\vartheta})}{V_{NWW}(x_i, \hat{\vartheta})} \right] \\ &\quad - \frac{\hat{\alpha} \hat{\xi}}{n} \sum_{i=1}^n \delta i \frac{x_i^{2\hat{\gamma}} K_{NWW}^{\hat{\xi}-1}(x_i, \hat{\vartheta}) M_{NWW}(x_i, \hat{\vartheta})}{V_{NWW}(x_i, \hat{\vartheta}) [1 - \exp(-\hat{\lambda} x_i^{\hat{\gamma}})]} \left[ 1 + \frac{(\hat{\xi} - 1) \exp(-\hat{\lambda} x_i^{\hat{\gamma}})}{K_{NWW}(x_i, \hat{\vartheta})} \right] + \\ &\quad \frac{1}{n} \sum_{i=1}^n \delta i \left[ \frac{x_i^{2\hat{\gamma}} \exp(-2\hat{\lambda} x_i^{\hat{\gamma}})}{[1 - \exp(-\hat{\lambda} x_i^{\hat{\gamma}})]^2} \left[ 1 - \hat{\alpha} \hat{\xi} (\hat{\xi} - 1) K_{NWW}^{\hat{\xi}-2}(x_i, \hat{\vartheta}) \right] \right], \end{aligned}$$

$$\begin{aligned}
 \hat{i}_{44} = & \frac{1}{n} \sum_{i=1}^n \delta i \left\{ \frac{-1}{\hat{\gamma}^2} - \hat{\lambda} \sum_{i \in F} x_i^{\hat{\gamma}} \log^2(x_i) + \hat{\lambda} (1 - \hat{\lambda} x_i^{\hat{\gamma}}) \left[ \hat{\alpha} \hat{\xi} \frac{K_{NWW}^{\hat{\xi}-1}(x_i, \hat{\vartheta}) - 1}{1 - \exp(-\hat{\lambda} x_i^{\hat{\gamma}})} + \frac{(\hat{\xi} - 1)}{D_{NWW}(x_i, \hat{\vartheta})} \right] [M_{NWW}(x_i, \hat{\vartheta}) \log(x_i)] \right\} \\
 & + \frac{\hat{\lambda}^2}{n} \sum_{i=1}^n \delta i \frac{M_{NWW}^2(x_i, \hat{\vartheta})}{[1 - \exp(-\hat{\lambda} x_i^{\hat{\gamma}})]^2} \left\{ 1 - \hat{\alpha} \hat{\xi} K_{NWW}^{\hat{\xi}-1}(x_i, \hat{\vartheta}) \left[ 1 + \frac{(\hat{\xi} - 1)}{K_{NWW}(x_i, \hat{\vartheta})} \right] \right\} \\
 & - \frac{\hat{\lambda}^2 (\hat{\xi} - 1)}{n} \sum_{i=1}^n \delta i \frac{M_{NWW}^2(x_i, \hat{\vartheta})}{D_{NWW}(x_i, \hat{\vartheta})} \left[ \frac{1}{1 - \exp(-\hat{\lambda} x_i^{\hat{\gamma}})} + \frac{1}{D_{NWW}(x_i, \hat{\vartheta})} \right] \\
 & + \frac{\hat{\alpha} \hat{\xi} \hat{\lambda}^2}{n} \sum_{i=1}^n \delta i \frac{x_i^{2\hat{\gamma}} \log^2(x_i) K_{NWW}^{\hat{\xi}-2}(x_i, \hat{\vartheta}) M_{NWW}(x_i, \hat{\vartheta}) \exp(-\hat{\lambda} x_i^{\hat{\gamma}})}{V_{NWW}(x_i, \hat{\vartheta}) [1 - \exp(-\hat{\lambda} x_i^{\hat{\gamma}})]} [\hat{\alpha} \hat{\xi} K_{NWW}^{\hat{\xi}}(x_i, \hat{\vartheta}) - (\hat{\xi} - 1)] \\
 & + \frac{\hat{\alpha} \hat{\xi} \hat{\lambda}}{n} \sum_{i=1}^n \delta i \frac{x_i^{\hat{\gamma}} \log^2(x_i) K_{NWW}^{\hat{\xi}-1}(x_i, \hat{\vartheta}) M_{NWW}(x_i, \hat{\vartheta})}{V_{NWW}(x_i, \hat{\vartheta})} \times \\
 & \left\{ 1 - \hat{\lambda} x_i^{\hat{\gamma}} \left[ \frac{1}{1 - \exp(-\hat{\lambda} x_i^{\hat{\gamma}})} - \hat{\alpha} \hat{\xi} \frac{K_{NWW}^{\hat{\xi}-1}(x_i, \hat{\vartheta}) Z_{NWW}(x_i, \hat{\vartheta})}{V_{NWW}(x_i, \hat{\vartheta})} \right] \right\},
 \end{aligned}$$

$$\begin{aligned}
 \hat{i}_{34} = & \frac{1}{n} \sum_{i=1}^n \delta i \left\{ M_{NWW}(x_i, \hat{\vartheta}) \left[ \frac{\hat{\alpha} \hat{\xi} K_{NWW}^{\hat{\xi}-1}(x_i, \hat{\vartheta}) - 1}{1 - \exp(-\hat{\lambda} x_i^{\hat{\gamma}})} + \frac{(\hat{\xi} - 1)}{D_{NWW}(x_i, \hat{\vartheta})} \right] (1 - \hat{\lambda} x_i^{\hat{\gamma}}) - [x_i^{\hat{\gamma}} \log(x_i)] \right\} \\
 & + \frac{\hat{\lambda}}{n} \sum_{i=1}^n \delta i \left\{ \frac{x_i^{2\hat{\gamma}} \log(x_i) \exp(-2\hat{\lambda} x_i^{\hat{\gamma}})}{[1 - \exp(-\hat{\lambda} x_i^{\hat{\gamma}})]^2} \left[ 1 - \hat{\alpha} \hat{\xi} K_{NWW}^{\hat{\xi}-1}(x_i, \hat{\vartheta}) - \hat{\alpha} (\hat{\xi} - 1) \hat{\xi} K_{NWW}^{\hat{\xi}-2}(x_i, \hat{\vartheta}) \right] \right\} \\
 & - \frac{\hat{\lambda} (\hat{\xi} - 1)}{n} \sum_{i=1}^n \delta i \frac{x_i^{2\hat{\gamma}} \log(x_i) \exp(-2\hat{\lambda} x_i^{\hat{\gamma}})}{D_{NWW}(x_i, \hat{\vartheta})} \left[ \frac{1}{D_{NWW}(x_i, \hat{\vartheta})} + \frac{1}{1 - \exp(-\hat{\lambda} x_i^{\hat{\gamma}})} \right] \\
 & + \frac{\hat{\alpha}^2 \hat{\xi}^2 \hat{\lambda}}{n} \sum_{i=1}^n \delta i \frac{x_i^{2\hat{\gamma}} \log(x_i) K_{NWW}^{2\hat{\xi}-2}(x_i, \hat{\vartheta}) M_{NWW}(x_i, \hat{\vartheta})}{V_{NWW}(x_i, \hat{\vartheta})} \left[ \frac{\exp(-\hat{\lambda} x_i^{\hat{\gamma}})}{1 - \exp(-\hat{\lambda} x_i^{\hat{\gamma}})} + \frac{M_{NWW}(x_i, \hat{\vartheta})}{V_{NWW}(x_i, \hat{\vartheta})} \right] \\
 & + \frac{\hat{\alpha} \hat{\xi}}{n} \sum_{i=1}^n \delta i \frac{x_i^{\hat{\gamma}} \log(x_i) K_{NWW}^{\hat{\xi}-1}(x_i, \hat{\vartheta}) M_{NWW}(x_i, \hat{\vartheta})}{V_{NWW}(x_i, \hat{\vartheta})} \left\{ 1 - \hat{\lambda} x_i^{\hat{\gamma}} \left[ \frac{1 + (\hat{\xi} - 1) \exp(-\hat{\lambda} x_i^{\hat{\gamma}}) K_{NWW}^{-1}(x_i, \hat{\vartheta})}{1 - \exp(-\hat{\lambda} x_i^{\hat{\gamma}})} \right] \right\},
 \end{aligned}$$

$$\begin{aligned}
 \hat{i}_{12} &= \frac{1}{n} \sum_{i=1}^n \delta_i \frac{\log[K_{NWW}(x_i, \hat{\vartheta})] K_{NWW}^{\hat{\xi}}(x_i, \hat{\vartheta})}{1 - \exp[-\hat{\alpha} K_{NWW}^{\hat{\xi}}(x_i, \hat{\vartheta})]} \left[ \frac{K_{NWW}^{\hat{\xi}}(x_i, \hat{\vartheta}) \exp[-\hat{\alpha} K_{NWW}^{\hat{\xi}}(x_i, \hat{\vartheta})]}{1 - \exp[-\hat{\alpha} K_{NWW}^{\hat{\xi}}(x_i, \hat{\vartheta})]} - 1 \right], \\
 \hat{i}_{13} &= \frac{\hat{\xi}}{n} \sum_{i=1}^n \delta_i \frac{x_i^{\hat{\gamma}} K_{NWW}^{\hat{\xi}-1}(x_i, \hat{\vartheta})}{1 - \exp(-\hat{\lambda} x_i^{\hat{\gamma}})} \left[ \exp(-\hat{\lambda} x_i^{\hat{\gamma}}) + \frac{M_{NWW}(x_i, \hat{\vartheta})}{1 - \exp[-\hat{\alpha} K_{NWW}^{\hat{\xi}}(x_i, \hat{\vartheta})]} \right] \\
 &\quad - \frac{\hat{\alpha} \hat{\xi}}{n} \sum_{i=1}^n \delta_i \frac{x_i^{\hat{\gamma}} K_{NWW}^{2\hat{\xi}-1}(x_i, \hat{\vartheta}) M_{NWW}(x_i, \hat{\vartheta})}{V_{NWW}(x_i, \hat{\vartheta})} \left[ \frac{1}{1 - \exp[-\hat{\alpha} K_{NWW}^{\hat{\xi}}(x_i, \hat{\vartheta})]} \right] \\
 \hat{i}_{14} &= \frac{\hat{\xi} \hat{\lambda}}{n} \sum_{i=1}^n \delta_i \frac{x_i^{\hat{\gamma}} \log(x_i) K_{NWW}^{\hat{\xi}-1}(x_i, \hat{\vartheta})}{1 - \exp(-\hat{\lambda} x_i^{\hat{\gamma}})} \left[ \exp(-\hat{\lambda} x_i^{\hat{\gamma}}) + \frac{M_{NWW}(x_i, \hat{\vartheta})}{1 - \exp[-\hat{\alpha} K_{NWW}^{\hat{\xi}}(x_i, \hat{\vartheta})]} \right] \\
 &\quad - \frac{\hat{\alpha} \hat{\xi} \hat{\lambda}}{n} \sum_{i=1}^n \delta_i \frac{x_i^{\hat{\gamma}} \log(x_i) K_{NWW}^{2\hat{\xi}-1}(x_i, \hat{\vartheta}) M_{NWW}(x_i, \hat{\vartheta})}{V_{NWW}(x_i, \hat{\vartheta})} \left[ \frac{1}{1 - \exp[-\hat{\alpha} K_{NWW}^{\hat{\xi}}(x_i, \hat{\vartheta})]} \right], \\
 \hat{i}_{23} &= \frac{1}{n} \sum_{i=1}^n \delta_i \frac{x_i^{\hat{\gamma}} \exp(-\hat{\lambda} x_i^{\hat{\gamma}})}{D_{NWW}(x_i, \hat{\vartheta})} - \frac{\hat{\alpha}^2 \hat{\xi}}{n} \sum_{i=1}^n \delta_i \frac{x_i^{\hat{\gamma}} \log[K_{NWW}(x_i, \hat{\vartheta})] K_{NWW}^{2\hat{\xi}-1}(x_i, \hat{\vartheta}) M_{NWW}(x_i, \hat{\vartheta})}{V_{NWW}(x_i, \hat{\vartheta}) [1 - \exp[-\hat{\alpha} K_{NWW}^{\hat{\xi}}(x_i, \hat{\vartheta})]]} \\
 &\quad + \frac{\hat{\alpha}}{n} \sum_{i=1}^n \delta_i \left\{ x_i^{\hat{\gamma}} K_{NWW}^{\hat{\xi}-1}(x_i, \hat{\vartheta}) \left[ \frac{\exp(-\hat{\lambda} x_i^{\hat{\gamma}})}{1 - \exp(-\hat{\lambda} x_i^{\hat{\gamma}})} + \frac{M_{NWW}(x_i, \hat{\vartheta})}{V_{NWW}(x_i, \hat{\vartheta})} \right] \{1 + \hat{\xi} \log[K_{NWW}(x_i, \hat{\vartheta})]\} \right\}, \\
 \hat{i}_{24} &= \frac{1}{n} \sum_{i=1}^n \delta_i \left[ \hat{\lambda} \frac{M_{NWW}(x_i, \hat{\vartheta})}{D_{NWW}(x_i, \hat{\vartheta})} \right] - \frac{\hat{\alpha}^2 \hat{\xi} \hat{\lambda}}{n} \sum_{i=1}^n \delta_i \frac{x_i^{\hat{\gamma}} \log(x_i) \log[K_{NWW}(x_i, \hat{\vartheta})] K_{NWW}^{2\hat{\xi}-1}(x_i, \hat{\vartheta}) M_{NWW}(x_i, \hat{\vartheta})}{V_{NWW}(x_i, \hat{\vartheta}) \{1 - \exp[-\hat{\alpha} K_{NWW}^{\hat{\xi}}(x_i, \hat{\vartheta})]\}} \\
 &\quad + \frac{\hat{\alpha} \hat{\lambda}}{n} \sum_{i \in F} \delta_i \left\{ K_{NWW}^{\hat{\xi}-1}(x_i, \hat{\vartheta}) \left[ \frac{M_{NWW}(x_i, \hat{\vartheta})}{1 - \exp(-\hat{\lambda} x_i^{\hat{\gamma}})} + \frac{x_i^{\hat{\gamma}} \log(x_i) M_{NWW}(x_i, \hat{\vartheta})}{V_{NWW}(x_i, \hat{\vartheta})} \right] \{1 + \hat{\xi} \log[K_{NWW}(x_i, \hat{\vartheta})]\} \right\}
 \end{aligned}$$

Où :

$$V_{NWW}(x, \vartheta) = [1 - \exp(-\lambda x^\gamma)] \times [1 - \exp(-\alpha \{-\log[1 - \exp(-\lambda x^\gamma)]\}^\xi)].$$

## 2.5 Simulation numérique

### Simulation des EMV dans le cas complet

À l'aide du logiciel *R*, nous proposons une étude par simulation numérique sur 12,000 échantillons de données complètes de la distribution N-Weibull-Weibull, avec les va-

leurs initiales ( $\alpha = 1, \xi = 1.5, \lambda = 2.5, \gamma = 1.3$ ). La méthode de *Newton-Raphson* disponible dans le package Barzilai-Borwein (*BB*) du logiciel statistique *R* est utilisée pour calculer les estimateurs du maximum de vraisemblance  $\hat{\vartheta} = (\hat{\alpha}, \hat{\xi}, \hat{\lambda}, \hat{\gamma})^T$  du modèle *NWW*, ainsi que leurs erreurs quadratiques moyennes (*MSE*) pour différentes tailles d'échantillon  $n = 20, 50, 100, 150, 250, 500, 750$ . Les résultats sont illustrés dans la Table 2.1

$N = 12,000$	$n = 20$	$n = 50$	$n = 100$	$n = 150$	$n = 250$	$n = 500$	$n = 750$
$\hat{\alpha}$	0.9894	1.0229	1.0018	1.0013	1.0012	1.0007	1.0005
<i>MSE</i>	$8.4764 \times 10^{-4}$	$6.9313 \times 10^{-4}$	$5.8781 \times 10^{-4}$	$5.0308 \times 10^{-5}$	$4.5366 \times 10^{-5}$	$4.1171 \times 10^{-5}$	$3.6648 \times 10^{-5}$
$\hat{\xi}$	1.5084	1.5070	1.5055	1.5034	1.4986	1.4977	1.4953
<i>MSE</i>	$3.7397 \times 10^{-4}$	$2.0336 \times 10^{-4}$	$1.2016 \times 10^{-4}$	$8.7691 \times 10^{-5}$	$6.5800 \times 10^{-5}$	$5.6558 \times 10^{-5}$	$4.3126 \times 10^{-5}$
$\hat{\lambda}$	2.5093	2.5100	2.5135	2.5041	2.4987	2.4970	2.4956
<i>MSE</i>	$7.9374 \times 10^{-4}$	$4.1065 \times 10^{-4}$	$2.5210 \times 10^{-4}$	$7.7870 \times 10^{-5}$	$1.3390 \times 10^{-5}$	$9.4950 \times 10^{-6}$	$7.6621 \times 10^{-6}$
$\hat{\gamma}$	1.3266	1.3235	1.3201	1.3183	1.3172	1.3154	1.3088
<i>MSE</i>	$7.5275 \times 10^{-4}$	$5.5352 \times 10^{-4}$	$4.0184 \times 10^{-4}$	$3.3086 \times 10^{-4}$	$2.9222 \times 10^{-4}$	$2.3623 \times 10^{-4}$	$9.6658 \times 10^{-5}$

TABLE 2.1 – Les valeurs moyennes des EMV des paramètres et leurs MSE pour les données complètes du modèle *NWW*.

Pour affirmer les résultats de la simulation, on trace dans la Figure 2.3 la courbe de la valeur absolue des EMV par rapport aux différentes tailles d'échantillon utilisées ( $n = 20, 50, 100, 150, 250, 500, 750$ )

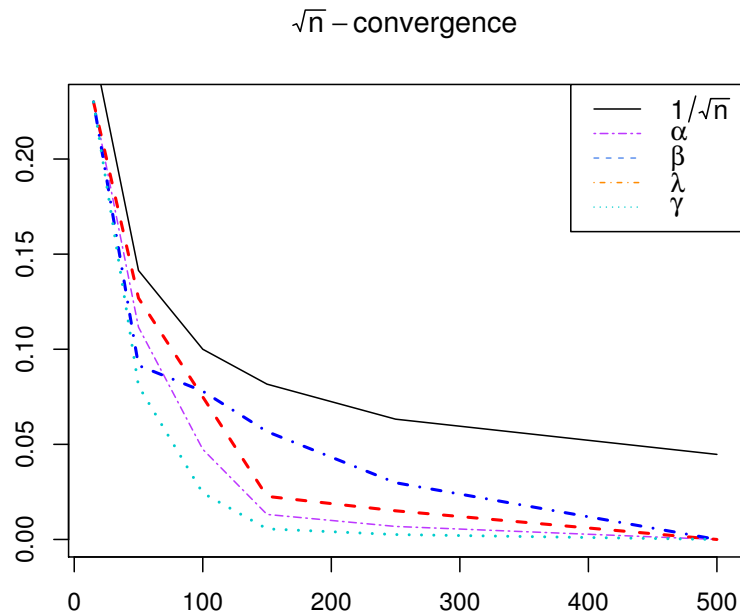


FIGURE 2.3 –  $\sqrt{n}$ -convergence des EMV de la distribution NWW.

À partir de la Figure 2.3 et les résultats dans Table 2.1, nous pouvons conclure que les EMV des paramètres obtenus  $\hat{\alpha}, \hat{\xi}, \hat{\lambda}, \hat{\gamma}$  de la distribution NWW sont  $\sqrt{n}$ -convergentes.

### Simulation des EMV censurés

Des données censurées à droite provenant de la distribution N-Weibull-Weibull ont été simulées 12,000 fois pour les valeurs initiales des paramètres, ( $\alpha = 9.1, \xi = 1.5, \lambda = 3.06, \gamma = 3.1$ ). En utilisant le package BB disponible dans le logiciel R, on a calculé  $\hat{\alpha}, \hat{\xi}, \hat{\lambda}, \hat{\gamma}$ , les estimateurs du maximum de vraisemblance des paramètres inconnus ainsi que leurs erreurs quadratiques moyennes (MSE) pour différentes tailles d'échantillons ( $n = 20, 50, 100, 150, 250, 500, 750$ ).

$N = 12,000$	$n = 20$	$n = 50$	$n = 100$	$n = 150$	$n = 250$	$n = 500$	$n = 750$
$\hat{\alpha}$	9.1838	9.1716	9.1563	9.1012	9.1004	9.1001	8.9875
MSE	$3.1764 \times 10^{-3}$	$8.2267 \times 10^{-4}$	$5.9015 \times 10^{-4}$	$3.1776 \times 10^{-4}$	$1.5389 \times 10^{-4}$	$7.6845 \times 10^{-5}$	$5.2144 \times 10^{-5}$
$\hat{\xi}$	1.5236	1.5133	1.5103	1.5047	1.5005	1.4999	1.4992
MSE	$3.9028 \times 10^{-3}$	$1.5987 \times 10^{-3}$	$6.6639 \times 10^{-4}$	$4.7225 \times 10^{-4}$	$1.0661 \times 10^{-4}$	$8.4932 \times 10^{-5}$	$6.2647 \times 10^{-5}$
$\hat{\lambda}$	3.0928	3.0856	3.0831	3.0705	3.0640	3.0611	3.0599
MSE	$8.3684 \times 10^{-3}$	$2.2975 \times 10^{-3}$	$1.2167 \times 10^{-3}$	$7.7233 \times 10^{-4}$	$3.221 \times 10^{-4}$	$1.0472 \times 10^{-4}$	$7.5438 \times 10^{-5}$
$\hat{\gamma}$	3.2308	3.1639	3.1344	3.1268	3.1036	3.1002	2.9967
MSE	$1.5534 \times 10^{-2}$	$3.1812 \times 10^{-3}$	$4.6265 \times 10^{-4}$	$2.1137 \times 10^{-4}$	$7.2128 \times 10^{-5}$	$3.9045 \times 10^{-5}$	$1.0871 \times 10^{-5}$

TABLE 2.2 – Les valeurs moyennes des EMV des paramètres et leurs MSE pour les données censurées du modèle *NWW*.

On remarque que les erreurs quadratiques moyennes sont très petites. Pour confirmer la convergence des estimateurs obtenus, on trace dans la Figure 2.4 la courbe de ceux-ci par rapport à la taille de l'échantillon correspondant.

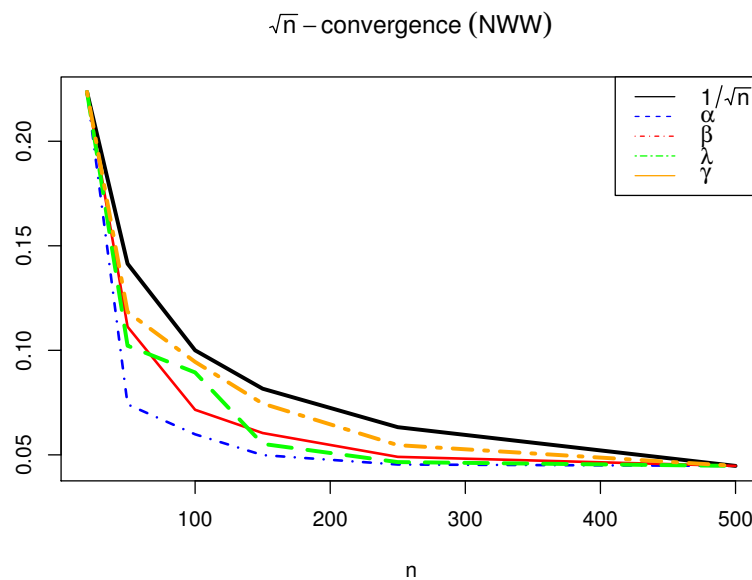


FIGURE 2.4 –  $\sqrt{n}$ -convergence des EMV pour les données censurées de la distribution *NWW*.

On peut affirmer que les estimateurs du maximum de vraisemblance sont  $\sqrt{n}$ -convergents.

### 3 Test d'ajustement pour le modèle *NWW*

#### 3.1 Critère du test d'ajustement dans le cas complet

Pour tester la validité d'un modèle statistique, différentes techniques peuvent être utilisées comme les statistiques d'Anderson-Darling, de Kolmogorov-Smirnov, Cramer-Von-Mises et beaucoup d'autres et particulièrement le test du chi-deux. Néanmoins celui-ci ne peut être utilisé quand les paramètres sont inconnus ou bien les données sont censurées. Nikulin (1973) a introduit un test du chi-deux modifié pour la famille des distributions continues avec des paramètres de forme et d'échelle inconnus et des données non-groupées. De leur côté Rao-Robson (1974) ont obtenu le même résultat pour les familles exponentielles. Depuis, cette statistique est connue sous le nom de la *NRR* (Nikulin-Rao-Robson) statistique.

Basée sur cette approche, on se propose de construire un critère de test du type du chi-deux modifié pour le modèle *NWW* dans le cas de paramètres inconnus.

Considérons l'échantillon  $X = (X_1, \dots, X_n)^T$ , supposons que l'hypothèse nulle  $H_0$  est vérifiée, c'est-à-dire que les taux de défaillance sont distribués selon la distribution *N-Weibull-Weibull* :

$$H_0 : P(X_i \leq x) = F_{NWW}(x, \vartheta), \quad \vartheta = (\alpha, \xi, \lambda, \gamma)^T$$

Les données initiales sont groupées en  $r$  intervalles  $I_j = (a_{j-1}, a_j]$ ,  $j = \overline{1, r}$ , les classes de regroupement sont supposées équiprobables, donc les probabilités correspondantes sont obtenues par :

$$p_j(\hat{\vartheta}_n) = \int_{a_{j-1}}^{a_j} dF(t, \hat{\vartheta}_n) = \frac{1}{r}, \quad j = \overline{1, r}.$$

Après avoir calculé les estimateurs du maximum de vraisemblance du vecteur de paramètres inconnus  $\hat{\vartheta} = (\hat{\alpha}, \hat{\xi}, \hat{\lambda}, \hat{\gamma})^T$ , les limites  $\hat{a}_j$  estimées des intervalles  $I_j$  sont obtenues comme suit :

$$\hat{a}_j = \left[ \frac{\log \left( 1 - \exp \left\{ - \left[ - \frac{\log(\frac{j}{r})}{\hat{\alpha}} \right]^{\frac{1}{\hat{\xi}}} \right\} \right)}{\hat{\lambda}} \right]^{\frac{1}{\hat{\gamma}}}.$$

Les fréquences empiriques  $\nu = (\nu_1, \dots, \nu_r)^T$  obtenues par le regroupement des données initiales dans les intervalles  $I_j$  sont obtenues comme suit :

$$\nu_j = \text{card} \{i : x_i \in I_j, \} i = \overline{1, n}.$$

### 3.1.1 Les éléments de la statistique du test

Pour construire une statistique  $Y^2$  (voir chapitre 1, section 3-1) pour le modèle *NWW* :

$$\begin{aligned} Y^2 &= X_n^2 + Q, \\ Q &= \frac{1}{n} L^T(\hat{\vartheta}) G^{-1} L(\hat{\vartheta}), \end{aligned}$$

on procède comme suit :

**Les éléments du vecteur  $\widehat{L}$  :**

On a :

$$p_j(\vartheta) = F(a_j) - F(a_{j-1}),$$

alors,

$$\begin{aligned} \frac{\partial p_j(\widehat{\vartheta})}{\partial \widehat{\alpha}} &= [K_{NWW}(\widehat{a}_j)]^{\widehat{\xi}} F_{NWW}(\widehat{a}_j) - [K_{NWW}(\widehat{a}_{j-1})]^{\widehat{\xi}} F_{NWW}(\widehat{a}_{j-1}), \\ \frac{\partial p_j(\widehat{\vartheta})}{\partial \widehat{\xi}} &= \widehat{\alpha} \log [K_{NWW}(\widehat{a}_j)] \times [K_{NWW}(\widehat{a}_j)]^{\widehat{\xi}} F_{NWW}(\widehat{a}_j) - \widehat{\alpha} \log [K_{NWW}(\widehat{a}_{j-1})] \\ &\quad \times [K_{NWW}(\widehat{a}_{j-1})]^{\widehat{\xi}} F_{NWW}(\widehat{a}_{j-1}), \\ \frac{\partial p_j(\widehat{\vartheta})}{\partial \widehat{\lambda}} &= \widehat{\alpha} \widehat{\xi} \left[ \frac{\widehat{a}_{j-1} [K_{NWW}(\widehat{a}_{j-1})]^{\widehat{\xi}-1} \times D_{NWW}(\widehat{a}_{j-1})}{1 - \exp(-\widehat{\lambda} \widehat{a}_{j-1}^{\widehat{\gamma}})} - \frac{\widehat{a}_j [K_{NWW}(\widehat{a}_j)]^{\widehat{\xi}-1} \times D_{NWW}(\widehat{a}_j)}{1 - \exp(-\widehat{\lambda} \widehat{a}_j^{\widehat{\gamma}})} \right], \\ \frac{\partial p_j(\widehat{\vartheta})}{\partial \widehat{\gamma}} &= \widehat{\alpha} \widehat{\xi} \widehat{\lambda} \left[ \frac{\widehat{a}_{j-1}^{\widehat{\gamma}} \log(\widehat{a}_{j-1}) [K_{NWW}(\widehat{a}_{j-1})]^{\widehat{\xi}-1} \times D_{NWW}(\widehat{a}_{j-1})}{1 - \exp(-\widehat{\lambda} \widehat{a}_{j-1}^{\widehat{\gamma}})} - \frac{\widehat{a}_j^{\widehat{\gamma}} \log(\widehat{a}_j) [K_{NWW}(\widehat{a}_j)]^{\widehat{\xi}-1} \times D_{NWW}(\widehat{a}_j)}{1 - \exp(-\widehat{\lambda} \widehat{a}_j^{\widehat{\gamma}})} \right], \end{aligned}$$

où :

$$\begin{aligned} K_{NWW}(x, \vartheta) &= \{-\log[1 - \exp(-\lambda x^\gamma)]\}, \\ D_{NWW}(x, \vartheta) &= \exp\{-\alpha \{-\log[1 - \exp(-\lambda x^\gamma)]\}^\xi - \lambda x^\gamma\}. \end{aligned}$$

Ainsi, le vecteur  $\widehat{L}_{NWW} = (\widehat{L}_1, \dots, \widehat{L}_s)^T$  est obtenu par :

$$\widehat{L}_{NWW}(\widehat{\vartheta}) = \left[ \widehat{L}_1(\widehat{\vartheta}) = \sum_{j=1}^r \frac{v_j}{p_j} \frac{\partial p_j(\widehat{\vartheta})}{\partial \widehat{\alpha}}, \widehat{L}_2(\widehat{\vartheta}) = \sum_{j=1}^r \frac{v_j}{p_j} \frac{\partial p_j(\widehat{\vartheta})}{\partial \widehat{\xi}}, \widehat{L}_3(\widehat{\vartheta}) = \sum_{j=1}^r \frac{v_j}{p_j} \frac{\partial p_j(\widehat{\vartheta})}{\partial \widehat{\lambda}}, \widehat{L}_4(\widehat{\vartheta}) = \sum_{j=1}^r \frac{v_j}{p_j} \frac{\partial p_j(\widehat{\vartheta})}{\partial \widehat{\gamma}} \right]^T.$$

Les éléments de la matrice  $\widehat{G}$  sont :

$$\begin{aligned}
 \widehat{g}_{11} &= \widehat{i}_{11} - \sum_{j=1}^r \frac{1}{p_j} \left( \frac{\partial p_j(\widehat{\vartheta})}{\partial \widehat{\alpha}} \right)^2, \quad \widehat{g}_{22} = \widehat{i}_{22} - \sum_{j=1}^r \frac{1}{p_j} \left( \frac{\partial p_j(\widehat{\vartheta})}{\partial \widehat{\xi}} \right)^2, \\
 \widehat{g}_{33} &= \widehat{i}_{33} - \sum_{j=1}^r \frac{1}{p_j} \left( \frac{\partial p_j(\widehat{\vartheta})}{\partial \widehat{\lambda}} \right)^2, \quad \widehat{g}_{44} = \widehat{i}_{44} - \sum_{j=1}^r \frac{1}{p_j} \left( \frac{\partial p_j(\widehat{\vartheta})}{\partial \widehat{\gamma}} \right)^2, \\
 \widehat{g}_{12} &= \widehat{i}_{12} - \sum_{j=1}^r \frac{1}{p_j} \left( \frac{\partial p_j(\widehat{\vartheta})}{\partial \widehat{\alpha}} \frac{\partial p_j(\widehat{\vartheta})}{\partial \widehat{\xi}} \right), \quad \widehat{g}_{13} = \widehat{i}_{13} - \sum_{j=1}^r \frac{1}{p_j} \left( \frac{\partial p_j(\widehat{\vartheta})}{\partial \widehat{\alpha}} \frac{\partial p_j(\widehat{\vartheta})}{\partial \widehat{\lambda}} \right), \\
 \widehat{g}_{14} &= \widehat{i}_{14} - \sum_{j=1}^r \frac{1}{p_j} \left( \frac{\partial p_j(\widehat{\vartheta})}{\partial \widehat{\alpha}} \frac{\partial p_j(\widehat{\vartheta})}{\partial \widehat{\gamma}} \right), \quad \widehat{g}_{23} = \widehat{i}_{23} - \sum_{j=1}^r \frac{1}{p_j} \left( \frac{\partial p_j(\widehat{\vartheta})}{\partial \widehat{\xi}} \frac{\partial p_j(\widehat{\vartheta})}{\partial \widehat{\lambda}} \right), \\
 \widehat{g}_{24} &= \widehat{i}_{24} - \sum_{j=1}^r \frac{1}{p_j} \left( \frac{\partial p_j(\widehat{\vartheta})}{\partial \widehat{\xi}} \frac{\partial p_j(\widehat{\vartheta})}{\partial \widehat{\gamma}} \right), \quad \widehat{g}_{34} = \widehat{i}_{34} - \sum_{j=1}^r \frac{1}{p_j} \left( \frac{\partial p_j(\widehat{\vartheta})}{\partial \widehat{\lambda}} \frac{\partial p_j(\widehat{\vartheta})}{\partial \widehat{\gamma}} \right), \quad j = \overline{1, r}.
 \end{aligned}$$

La statistique  $Y^2$  obtenue suit une distribution du chi-deux ( $\chi^2$ ) à  $(r - 1)$  degrés de liberté.

### 3.1.2 Simulation des valeurs du critère de test

Pour tester l'hypothèse nulle  $H_0$  selon laquelle un échantillon de données provient de la distribution N-Weibull-Weibul (*NWW*) et pour montrer les performances de la statistique du test *NRR*, nous proposons une étude par simulation numérique. On génère 12,000 échantillons à partir de la distribution *NWW* avec différentes tailles  $n = 20, 50, 100, 150, 250, 500, 750$ . Nous calculons les valeurs de  $Y^2$  et les comparons à leurs valeurs théoriques correspondantes ( $\chi_c^2$ ) pour différents seuils de signification,  $\epsilon = 1\%, 5\%, 10\%$ . Les résultats sont affichés dans la Table 2.3

		<i>N</i> = 12,000						
$\epsilon \backslash n$		20	50	100	150	250	500	750
1%		0.0094	0.0136	0.0098	0.0088	0.0122	0.0108	0.0112
5%		0.0342	0.0414	0.0426	0.0442	0.0497	0.0505	0.0509
10%		0.0578	0.0772	0.0806	0.0878	0.0922	0.1006	0.1056

TABLE 2.3 – Comparaison des risques théoriques et des risques empiriques pour les données complètes du modèle *NWW*.

Les valeurs des niveaux de signification empiriques de  $Y^2$  sont assez proches de celles des niveaux théoriques correspondants de la distribution du chi-deux  $\chi^2$  à  $(r-1)$  degrés de liberté, ce qui nous permet de conclure que le test statistique  $Y^2$  construit dans ce travail est apte à ajuster des ensembles de données complètes à la distribution *NWW*.

### 3.2 Critère du test d'ajustement dans le cas censuré

Pour valider le modèle N-Weibull-Weibull, nous utilisons l'approche proposée par Bagdonavicius et Nikulin (2011), détaillée dans le chapitre 1, pour construire un critère de test capable de vérifier si une série de données censurées à droite est modélisée par la distribution *NWW*. Ce test est basé sur l'estimation du maximum de vraisemblance sur des données non-groupées.

Soit  $X_i = (T_i \wedge C_i)$  un échantillon censuré à droite. Supposons que l'hypothèse nulle  $H_0$  est vraie, telle que  $X_i$  suit la distribution N-Weibull Weibull. Les données observées sont regroupées en  $k$  intervalles  $I_j$  :

$$I_j = (a_{j-1}, a_j], j = \overline{1, k}, \quad a_0 = 0, \quad a_k = \max(X_{(l)}, \tau).$$

Le choix des limites  $\hat{a}_j$  des intervalles de regroupement  $I_j$  est obtenu comme suit :

$$\hat{a}_j = \left( -\frac{1}{\hat{\lambda}} \log \left\{ 1 - \exp \left[ - \left( -\frac{1}{\hat{\alpha}} \log \left\{ 1 - \exp \left[ \frac{\sum_{l=1}^{i-1} H(X_{(l)}, \hat{\vartheta}) - E_j}{N - i + 1} \right] \right\} \right]^{\frac{1}{\hat{\xi}}} \right] \right\} \right)^{\frac{1}{\hat{\gamma}}}, j = \overline{1, k}$$

où,  $E_j = (N - i + 1) H(\hat{a}_j, \hat{\vartheta}) + \sum_{l=1}^{i-1} H(X_{(l)}, \hat{\vartheta})$ . La statistique  $Y_n^2$  est donnée par :

$$Y_n^2 = \sum_{j=1}^k \frac{(U_j - e_j)^2}{U_j} + Q,$$

tel que  $U_j$  est le nombre des défaillances observées et  $e_j$  le nombre de défaillances théoriques :

$$U_j = \sum_{i: X_i \in I_j} \delta_i,$$

$$e_j = \frac{E_k}{k}, \quad E_k = \sum_{i=1}^n H(x_i, \hat{\vartheta}).$$

### 3.2.1 Les éléments de la statistique du test

Pour le calcul de la statistique du test construite plus haut, nous devons déterminer toutes ses composantes qui sont la matrice  $\hat{C}$  et le vecteur  $W$ .

**Les éléments de matrice  $\hat{C} = (\hat{C})_{4 \times k}$  :**

$$\hat{C}_{1j} = \frac{1}{n} \sum_{i: X_i \in I_j} \delta_i \left[ \frac{1}{\hat{\alpha}} - \frac{K_{NWW}^{\hat{\xi}}(x_i, \hat{\vartheta})}{1 - \exp[-\hat{\alpha} K_{NWW}^{\hat{\xi}}(x_i, \hat{\vartheta})]} \right],$$

$$\hat{C}_{2j} = \frac{1}{n} \sum_{i: X_i \in I_j} \delta_i \left[ \frac{1}{\hat{\xi}} + \log[K_{NWW}(x_i, \hat{\vartheta})] - \hat{\alpha} \frac{\log[K_{NWW}(x_i, \hat{\vartheta})] K_{NWW}^{\hat{\xi}}(x_i, \hat{\vartheta})}{1 - \exp[-\hat{\alpha} K_{NWW}^{\hat{\xi}}(x_i, \hat{\vartheta})]} \right],$$

$$\begin{aligned}\widehat{C}_{3j} &= \frac{1}{n} \sum_{i: X_i \in I_j} \delta_i \left[ \frac{1}{\widehat{\lambda}} - x_i^{\widehat{\gamma}} + \widehat{\alpha} \widehat{\xi} x_i^{\widehat{\gamma}} K_{NWW}^{\widehat{\xi}-1}(x_i, \widehat{\vartheta}) Z_{NWW}(x_i, \widehat{\vartheta}) \right] \\ &\quad + \frac{1}{n} \sum_{i: X_i \in I_j} \delta_i \left[ + \frac{x_i^{\widehat{\gamma}} \exp(-\widehat{\lambda} x_i^{\widehat{\gamma}})}{1 - \exp(-\widehat{\lambda} x_i^{\widehat{\gamma}})} [\widehat{\alpha} \widehat{\xi} K_{NWW}^{\widehat{\xi}-1}(x_i, \widehat{\vartheta}) - 1] + (\widehat{\xi} - 1) \frac{x_i^{\widehat{\gamma}} \exp(-\widehat{\lambda} x_i^{\widehat{\gamma}})}{D_{NWW}(x_i, \widehat{\vartheta})} \right]\end{aligned}$$

$$\begin{aligned}\widehat{C}_{4j} &= \frac{1}{n} \sum_{i: X_i \in I_j} \delta_i \left[ \frac{1}{\widehat{\gamma}} + \widehat{\lambda} (\widehat{\xi} - 1) \frac{M_{NWW}(x_i, \widehat{\vartheta})}{D_{NWW}(x_i, \widehat{\vartheta})} + \widehat{\lambda} \frac{M_{NWW}(x_i, \widehat{\vartheta})}{1 - \exp(-\widehat{\lambda} x_i^{\widehat{\gamma}})} [\widehat{\alpha} \widehat{\xi} K_{NWW}^{\widehat{\xi}-1}(x_i, \widehat{\vartheta}) - 1] \right] \\ &\quad + \sum_{i: X_i \in I_j} \delta_i \left[ \log(x_i) [1 - \widehat{\lambda} x_i^{\widehat{\gamma}}] + \widehat{\alpha} \widehat{\xi} \widehat{\lambda} \frac{x_i^{\widehat{\gamma}} \log(x_i) K_{NWW}^{\widehat{\xi}-1}(x_i, \widehat{\vartheta}) Z_{NWW}(x_i, \widehat{\vartheta})}{[1 - \exp(-\widehat{\lambda} x_i^{\widehat{\gamma}})] \{1 - \exp[-\widehat{\alpha} K_{NWW}^{\widehat{\xi}}(x_i, \widehat{\vartheta})]\}} \right].\end{aligned}$$

Où :

$$\begin{aligned}D_{NWW}(x, \vartheta) &= \exp\left(-\alpha \{-\log[1 - \exp(-\lambda x^\gamma)]\}^\xi - \lambda x^\gamma\right) \\ K_{NWW}(x, \vartheta) &= \{-\log[1 - \exp(-\lambda x^\gamma)]\}, \\ M_{NWW}(x, \vartheta) &= x^\gamma \times \log(x) \times \exp(-\lambda x^\gamma), \\ Z_{NWW}(x, \vartheta) &= [1 - \exp(-\lambda x^\gamma)] \times \log[1 - \exp(-\lambda x^\gamma)], \\ V_{NWW}(x, \vartheta) &= [1 - \exp(-\lambda x^\gamma)] \times [1 - \exp(-\alpha \{-\log[1 - \exp(-\lambda x^\gamma)]\}^\xi)].\end{aligned}$$

On a :

$$\begin{aligned}W &= (W_1, \dots, W_s)^T = \widehat{C} \widehat{A}^{-1} Z, \quad \widehat{A}_j = \frac{U_j}{n}, \\ Z_j &= \frac{1}{\sqrt{n}} (U_j - e_j), \quad j = \overline{1, k}, \quad l, l' = 1, 2, 3, 4 \text{ et } \widehat{\vartheta} = (\widehat{\alpha}, \widehat{\xi}, \widehat{\lambda}, \widehat{\gamma}).\end{aligned}$$

Nous obtenons la statistique du test  $Y_n^2$  :

$$Y_n^2 = \sum_{i=1}^n \frac{(U_j - e_j)^2}{U_j} + W^T \times \left[ \widehat{i}_{ll'} - \sum_{j=1}^k \widehat{C}_{lj} \widehat{C}_{l'j} \widehat{A}_j^{-1} \right]^{-1} \times W,$$

cette statistique ( $Y_n^2$ ) suit une distribution de chi-deux à ( $k$ ) degrés de liberté.

### 3.2.2 Simulation des valeurs du critère de test

À l'aide du logiciel statistique *R*, on calcule la valeur de la statistique  $Y_n^2$  et on la compare au risque d'erreur théorique correspondant  $\chi_\epsilon^2$  pour différents niveaux de signification,  $\epsilon = 1\%, 5\%, 10\%$ .

		$N = 12,000$						
$\epsilon \backslash n$		20	50	100	150	250	500	750
1%		0.0087	0.0090	0.0096	0.0102	0.0108	0.0112	0.0115
5%		0.0326	0.0413	0.0442	0.0446	0.0486	0.0503	0.0509
10%		0.0853	0.0892	0.0964	0.1031	0.1046	0.1083	0.0110

TABLE 2.4 – Comparaison des risques théoriques et des risques empiriques pour les données censurées du modèle *NWW*.

De la Table 2.4 et en tenant compte des erreurs de simulation, on observe que les niveaux de signification empiriques  $Y_n^2$  sont assez proches des valeurs théoriques correspondantes de la distribution du chi-deux  $\chi_\epsilon^2$  avec  $k$  degrés de liberté. Ce qui nous permet de conclure que le test statistique construit peut ajuster adéquatement des données censurées à droite du modèle N-Weibull-Weibull (*NWW*).

## 4 Tests classiques

Dans cette section, des illustrations numériques sont présentées pour montrer les capacités de la distribution N-Weibull-Weibull via l'étude par simulation des critères de tests bien connus dans la littérature, Akaike Information Criteria (*AIC*), Consistent Akaike Information Criteria (*CAIC*), Bayesian Information Criteria (*BIC*), Hanan et Quinn Information Criteria (*HQIC*) et Kolmogorov Smirnov Criteria (*KS*). Nous comparons la

distribution  $NWW$  à quatre distributions alternatives : Weibull ( $W$ ), Inverse Weibull ( $IW$ ), Topp-Leone Weibull-Weibull ( $TLWW$ ) et Weibull-Weibull ( $WW$ ).

Nous avons réalisé 12,000 simulations de données issues des distributions :  $W$ ,  $IW$ ,  $TLWW$ ,  $WW$  et  $NWW$ , respectivement, pour différentes tailles d'échantillons et on a calculé les critères de test mentionnés ci-dessus. Les résultats sont illustrés dans la Table 2.5 :

$n$	Model	$AIC$	$BIC$	$CAIC$	$HQIC$	$KS$
10	$W$	31.25	32.06	31.22	31.64	31.18
	$IW$	30.66	30.98	31.57	31.41	31.83
	$WW$	31.87	31.33	32.14	31.05	30.8
	$TLWW$	31.04	31.07	31.11	31.28	31.22
	<b><math>NWW</math></b>	<b>30.21</b>	<b>30.45</b>	<b>30.77</b>	<b>30.13</b>	<b>31.04</b>
20	$W$	27.32	27.64	27.44	27.39	24.91
	$IW$	20.78	21.2	20.99	20.86	24.96
	$WW$	20.48	21.11	20.90	20.6	24.8
	$TLWW$	20.1	20.74	20.53	20.23	24.79
	<b><math>NWW</math></b>	<b>20.04</b>	<b>20.67</b>	<b>20.46</b>	<b>20.16</b>	<b>23.58</b>
30	$W$	25.13	26.88	26.91	27.03	25.89
	$IW$	22.64	25.46	20.08	19.44	24.72
	$WW$	22.18	23.31	20.31	20.25	25.01
	$TLWW$	21.95	23.11	21.22	19.66	24.64
	<b><math>NWW</math></b>	<b>20.22</b>	<b>20.82</b>	<b>19.85</b>	<b>19.37</b>	<b>24.05</b>
50	$W$	20.14	20.45	20.16	20.26	25.37
	$IW$	19.72	20.17	19.76	19.89	25.20
	$WW$	20.35	21.12	20.44	20.64	25.21
	$TLWW$	20.54	21.15	20.61	20.77	25.31
	<b><math>NWW</math></b>	<b>18.53</b>	<b>19.29</b>	<b>18.61</b>	<b>18.82</b>	<b>25.08</b>
100	$W$	20.2	20.41	20.21	20.28	25.48
	$IW$	21.18	21.49	21.31	20.56	25.44
	$WW$	20.31	20.93	20.33	20.56	25.39
	$TLWW$	21.56	20.93	21.17	21.32	25.44
	<b><math>NWW</math></b>	<b>18.65</b>	<b>19.28</b>	<b>18.68</b>	<b>18.91</b>	<b>25.38</b>

TABLE 2.5 – Valeurs des critères d’information  $AIC/BIC/CAIC/HQIC/KS$  pour les données simulées.

À partir de la Table 2.5, nous observons que la distribution  $NWW$  a les plus petites valeurs des critères de tests utilisés, ce qui nous amène à déduire que le modèle N-Weibull-Weibull est une bonne alternative à tous ces modèles.

## 5 Applications aux données réelles

Afin de démontrer l'applicabilité et l'utilité de cette nouvelle distribution N-Weibull-Weibull, nous utiliserons des données réelles complètes et censurées de différents domaines d'application.

### 5.1 Données économiques

Considérons ces 50 observations économiques disponibles dans "<https://data.world/datasets/data>"; qui représentent l'inflation annuelle en Algérie entre 1961 et 2011. On veut tester si elles peuvent être modélisées par la distribution N-Weibull-Weibull.

3.471720042, 2.351279502, 0.5493313109, 1.695183177, 1.501331249, 1.817814694, 1.312040991, 3.142056013, 1.921084494, 4.940445977, 17.15196386, 4.606461074, 9.627611595, 48.89659052, 5.914022113, 10.8405927, 11.92709947, 10.08512101, 13.98783797, 25.86203875, 14.35399954, 1.93979417, 6.804795897, 8.433505552, 4.972526406, 2.405343265, 8.842020401, 9.060963496, 16.01137351, 30.25959848, 53.78860423, 21.92611451, 13.62442466, 29.07764734, 28.5770375, 24.02190406, 7.001963063, 3.131088695, 10.85640762, 24.59809885, 0.7112095185, 1.906328849, 8.323802693, 10.62932921, 16.45925846, 11.28281156, 7.331055187, 14.60217944, 11.2666112, 16.24561679, 11.43116826.
---

TABLE 2.6 – Données d'inflation annuelle en Algérie (1961 – 2011).

#### - Analyse graphique :

Nous traçons dans la Figure 2.5 la fonction de répartition et la densité de probabilité estimée, ainsi que les QQ-plot et PP-plot pour comparer visuellement les quantiles et les probabilités théoriques de la distribution *NWW* et ceux de la distribution empirique

du taux d'inflation.

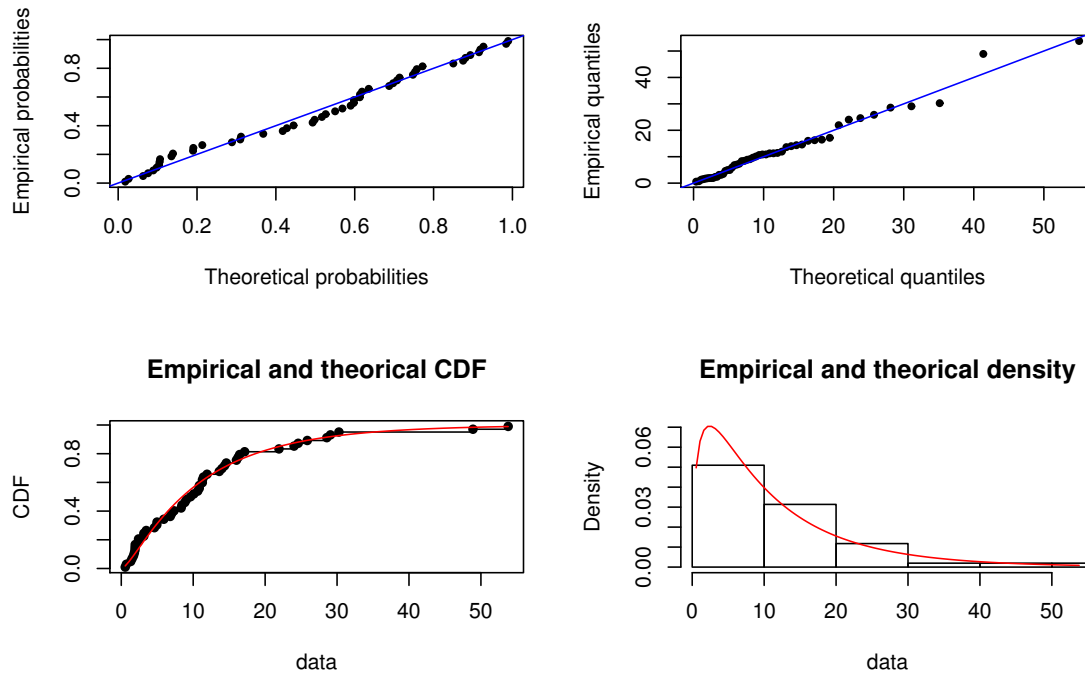


FIGURE 2.5 – pdf, cdf, QQ et PP plot de l'inflation annuelle en Algérie (1961 – 2011).

La Figure 2.5 nous permet de voir au premier aspect que cet ensemble de données peut être ajusté par la distribution *NWW*.

#### - Tests classiques :

Dans la Table 2.7 on compare la distribution *NWW* avec les distributions utilisées généralement dans la modélisation des données économiques, telles que *Gumbel*-exponentielle (*GE*), *Pareto* (*Pa*) et *Rayleigh* (*R*). Pour cela, nous calculons les statistiques des critères d'information *AIC*, *BIC*, *CAIC*, *HQIC* et *KS* pour ces données :

Model	<i>AIC</i>	<i>BIC</i>	<i>CAIC</i>	<i>HQIC</i>	<i>KS</i>
<b>NWW</b>	<b>361.63</b>	<b>369.36</b>	<b>362.50</b>	<b>364.58</b>	<b>0.0577</b>
<i>WW</i>	363.09	370.81	363.96	366.04	0.0733
<i>Rayleigh</i>	398.17	400.10	398.25	398.91	0.2633
<i>Pareto</i>	363.27	371.44	364.12	365.87	0.1157
<i>GE</i>	365.03	368.89	365.28	366.50	0.4548

TABLE 2.7 – Valeurs des critères d’information *AIC*, *BIC*, *CAIC*, *HQIC* et *KS*.

La Table 2.7 indique que le modèle *NWW* a les valeurs des critères d’information les plus faibles, ce qui nous permet de dire que le modèle *NWW* s’adapte mieux que les distributions alternatives à ces observations.

**- Le test NRR :**

Nous supposons vraie l’hypothèse nulle  $H_0$  que les données d’inflation utilisées sont ajustées par le modèle N-Weibull-Weibull. Nous utilisons la méthode du maximum de vraisemblance pour estimer le vecteur des paramètres de la distribution *NWW* :

$$\hat{\theta} = (\hat{\alpha}, \hat{\xi}, \hat{\lambda}, \hat{\gamma})^T = (0.7633, 2.7023, 0.0954, 0.8217)^T$$

Nous choisissons  $r = 5$  le nombre de classes de regroupement ( $I_j$ ) des données, nous estimons les bornes de ces classes  $\hat{a}_j$  et calculons les valeurs de  $\hat{v}_j$  et  $p_j$ , les résultats sont donnés dans la Table 2.8 :

$\hat{a}_j$	6.5450	9.7173	13.8092	20.9757	153.7886
$\hat{v}_j$	18	8	9	7	9
$p_j$	0.2	0.2	0.2	0.2	0.2

TABLE 2.8 – Valeurs de  $\hat{a}_j$ ,  $\hat{v}_j$  et  $p_j$ .

Par conséquent, nous obtenons la valeur de la statistique du test  $NRR$ ,  $Y^2 = 7.6422$ , puis nous la comparons à la valeur du chi-deux  $\chi^2$  à  $(r - 1)$  degrés de liberté pour différents niveaux de signification  $\epsilon = 1\%, 5\%, 10\%$  :

$$Y^2 < \chi_{1\%}^2(4) = 15.0862 \quad ; \quad Y^2 < \chi_{5\%}^2(4) = 11.0705 \quad ; \quad Y^2 < \chi_{10\%}^2(4) = 9.2363.$$

D'après les résultats obtenus, nous pouvons conclure que  $H_0$  n'est pas rejetée, ce qui nous amène à dire que ces données économiques peuvent être ajustées par le modèle N-Weibull-Weibull ( $NWW$ ) à différents risques d'erreur  $\epsilon$ .

## 5.2 Données de fiabilité

Notre objectif est d'étudier l'hypothèse nulle  $H_0$  selon laquelle la distribution  $NWW$  ajuste le temps de remplacement de sièges de soupapes pour 41 moteurs diesel. Cet ensemble de données rapporté par Meeker et al., (1998) présenté dans la Table 2.9 est disponible dans le package 'Survival' du logiciel statistique  $R$ .

<p>761*, 759*, 98, 667*, 326, 653, 653, 667*, 665*, 84, 667*, 87, 663*, 646, 653*, 92, 653*,  651*, 258, 328, 377, 621, 650*, 61, 539, 648*, 254, 276, 298, 640, 644*,  76, 538, 642*, 635, 641*, 349, 404, 561, 649*, 631*, 596*, 120, 479, 614*, 323,  449, 582*, 139, 139, 589*, 593*, 573, 589*, 165, 408, 604, 606*, 249, 594*,  344, 497, 613*, 265, 586, 595*, 166, 206, 348, 389*, 601*, 410, 581, 601*, 611*,  608*, 587*, 36, 603*, 202, 563, 570, 585*, 587*, 578*, 578*, 586*, 585*, 582*.</p>
--

TABLE 2.9 – Le temps de défaillances de 41 moteurs diesel.

\* représente la censure.

### - Analyse graphique :

En utilisant ces données, nous traçons dans la Figure 2.6, la courbe de la densité de probabilité et la courbe de la fonction de répartition correspondantes à la distribution

*NWW* ainsi que le PP-plot et le QQ-plot en les comparant à la distribution des données utilisées.

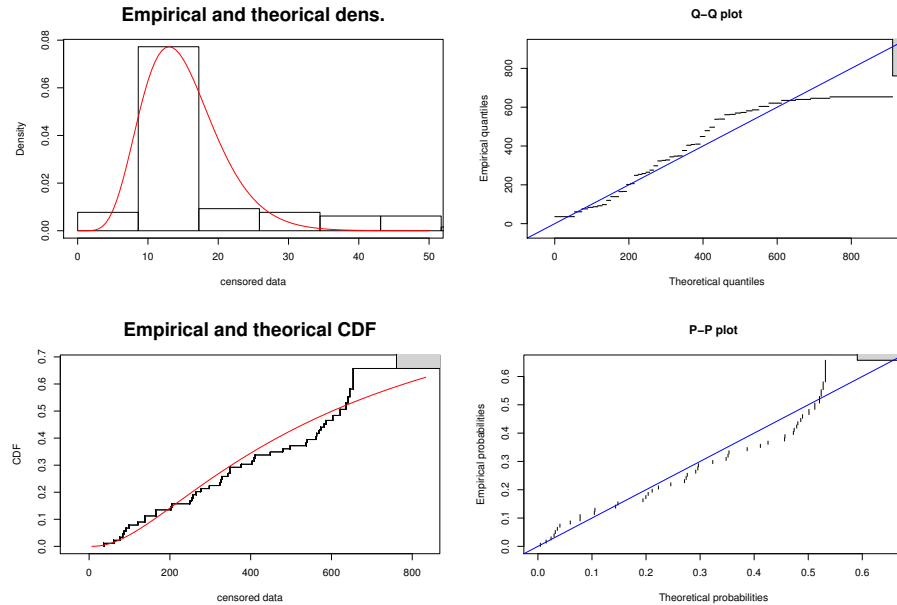


FIGURE 2.6 – pdf, cdf, QQ et PP plot du temps de défaillance de 40 moteurs diesel.

À partir de la Figure 2.6, nous pouvons déduire que la distribution N-Weibull-Weibull représente correctement ces observations.

**- Tests classiques :**

Ces données sont utilisées pour calculer les différents critères d'information classiques en vue de montrer la flexibilité du nouveau modèle *NWW*. Nous comparons dans la Table 2.10, la distribution N-Weibull-Weibull à des distributions communément utilisées dans la modélisation des données de fiabilité, Weibull-Weibull (*WW*), topp-leone-Weibull-Weibull (*TLWW*) et inverse-Weibull (*IW*) :

Model	<i>AIC</i>	<i>BIC</i>	<i>CAIC</i>	<i>HQIC</i>	<i>KS</i>
<b>NWW</b>	<b>1011.80</b>	<b>1021.75</b>	<b>1012.27</b>	<b>1015.81</b>	<b>0.1960</b>
<i>WW</i>	1177.93	1187.88	1178.41	1181.94	0.2140
<i>TLWW</i>	1240.29	1250.25	1240.77	1244.31	0.2832
<i>IW</i>	1306.29	1313.76	1306.58	1309.30	0.2862

TABLE 2.10 – Valeurs des critères d’information *AIC*, *BIC*, *CAIC*, *HQIC* et *KS*.

Les résultats obtenus confirment l’hypothèse nulle  $H_0$ .

**- Le test du chi-deux modifié :**

Étant donné que ces données sont censurées à droite, on utilise la statistique  $Y_n^2$  proposée dans ce travail pour vérifier  $H_0$ . À l’aide du logiciel statistique *R*, nous calculons les EMV de la distribution *NWW* :

$$\hat{\theta} = (\hat{\alpha}, \hat{\xi}, \hat{\lambda}, \hat{\gamma})^T = (19.6, 0.8425, 1.0005, 0.2)^T$$

On choisit  $r = 6$  intervalles de regroupement ( $I_j$ ) de ces observations, les éléments de la statistique  $Y_n^2$  du test de chi-deux modifié sont présentés dans la Table 2.11 :

$\hat{a}_j$	76.1399	145.6574	226.6525	323.7033	452.215	761.00
$\hat{U}_j$	2	7	4	7	11	17
$e_j$	5.8739	5.8739	5.8739	5.8739	5.8739	5.8739
$\hat{C}_{1j}$	-0.0013	-0.0031	-0.0009	-0.0007	-0.0003	0.0010
$\hat{C}_{2j}$	0.13157	0.4127	0.2031	0.3167	0.4541	0.6118
$\hat{C}_{3j}$	0.1147	0.3550	0.1724	0.2668	0.3811	0.5113
$\hat{C}_{4j}$	0.5964	2.0463	1.1234	1.8911	2.8734	4.2114

TABLE 2.11 – Valeurs de  $\hat{a}_j$ ,  $\hat{U}_j$ ,  $e_j$ ,  $\hat{C}_{1j}$ ,  $\hat{C}_{2j}$ ,  $\hat{C}_{3j}$  et  $\hat{C}_{4j}$ .

Alors, la valeur du critère de test obtenue est  $Y_n^2 = 6.4797$ . On la compare avec la statis-

tique du test du chi-deux  $\chi^2$  à  $k$  degrés de liberté pour différents seuils de signification  $\epsilon = 1\%$ ,  $\epsilon = 5\%$ ,  $\epsilon = 10\%$ , on a :

$$Y_n^2 < \chi_{1\%}^2(6) = 16.8118 \quad ; \quad Y_n^2 < \chi_{5\%}^2(6) = 12.5915 \quad ; \quad Y_n^2 < \chi_{10\%}^2(6) = 10.6446.$$

Les résultats obtenus affirment que cet ensemble de données peut parfaitement être modélisé par la distribution New-Weibull-Weibull (*NWW*).

À l'issue de cette étude du modèle N-Weibull-Weibull, nous pouvons affirmer que le modèle étudié est capable de modéliser adéquatement différents type de données provenant de différents domaines d'applications.

# Chapitre 3

## Le modèle N-Weibull-Rayleigh (NWR)

La distribution de *Rayleigh* est une distribution de probabilité continue nommée d'après le mathématicien anglais Lord Rayleigh en 1900, elle est particulièrement utile pour modéliser des variables aléatoires à valeurs positives résultant de la combinaison de variables gaussiennes indépendantes et distribuées de manière identique. Elle est appliquée dans divers domaines, notamment l'ingénierie, la physique, la sociologie et les télécommunications. Elle est couramment utilisée pour modéliser des variables telles que la vitesse du vent, la hauteur des vagues et l'ampleur des événements sismiques. Il est important de noter que même si la distribution de Rayleigh classique a ses avantages, elle n'est pas toujours la mieux adaptée à chaque ensemble de données. C'est ce qui nous a motivé à introduire la distribution de Rayleigh à la Nouvelle famille Weibull-G

Dans le présent chapitre, on introduit un nouveau modèle généralisé New-Weibull-Rayleigh, noté *NWR*, dont la distribution de base est une distribution de Rayleigh. Après le calcul des paramètres inconnus et des éléments de la matrice d'information de Fisher, on se propose de produire des critères de tests d'ajustement pour ce modèle dans le cas de données complètes et le cas de données censurées à droite. Des appli-

cations aux données réelles ont mis en valeur l'utilité de ce nouveau modèle à décrire différents phénomènes.

## 1 Présentation du modèle *NWR*

Soit  $X$  une variable aléatoire suivant une distribution de Rayleigh de paramètre  $\lambda$ ,

$$G(x) = 1 - \exp(-\lambda x^2)$$

En remplaçant la distribution de base  $G$  de la famille New-Weibull-G par celle de Rayleigh, on obtient un nouveau modèle appelé New-Weibull-Rayleigh (*NWR*) dont la fonction de répartition est donnée par :

$$F_{NWR}(x, \varsigma) = \exp\left\{-\alpha \left[-\log(1 - \exp(-\lambda x^2))\right]^\xi\right\}, x > 0, \varsigma = (\alpha, \xi, \lambda) > 0,$$

Alors, la densité de probabilité correspondante est :

$$\begin{aligned} f_{NWR}(x, \varsigma) &= \alpha \xi \frac{2\lambda x \exp(-\lambda x^2)}{1 - \exp(-\lambda x^2)} \left\{-\log[1 - \exp(-\lambda x^2)]\right\}^{\xi-1} \\ &\quad \times \exp\left(-\alpha \left\{-\log[1 - \exp(-\lambda x^2)]\right\}^\xi\right), \end{aligned}$$

On peut déduire, les fonctions de survie, du taux de hasard et du taux de hasard cumulé :

$$S_{NWR}(x, \varsigma) = 1 - \exp\left\{-\alpha \left[-\log(1 - \exp(-\lambda x^2))\right]^\xi\right\},$$

$$h_{NWR}(x, \varsigma) = \frac{\alpha \xi \frac{2\lambda x \exp(-\lambda x^2)}{1 - \exp(-\lambda x^2)} \{-\log[1 - \exp(-\lambda x^2)]\}^{\xi-1}}{1 - \exp\left(-\alpha \{-\log[1 - \exp(-\lambda x^2)]\}^{\xi}\right)} \times \exp\left(-\alpha \{-\log[1 - \exp(-\lambda x^2)]\}^{\xi}\right),$$

$$H_{NWR}(x, \varsigma) = -\log\left\{1 - \exp\left(-\alpha \left[-\log(1 - \exp(-\lambda x^2))\right]^{\xi}\right)\right\}.$$

Nous traçons dans la Figure 3.1 et la Figure 3.2 la courbe de la densité de probabilité et du taux de hasard de la distribution *NWR* et celle de la distribution *WR* dans le but de montrer la souplesse et la flexibilité de ce nouveau modèle.

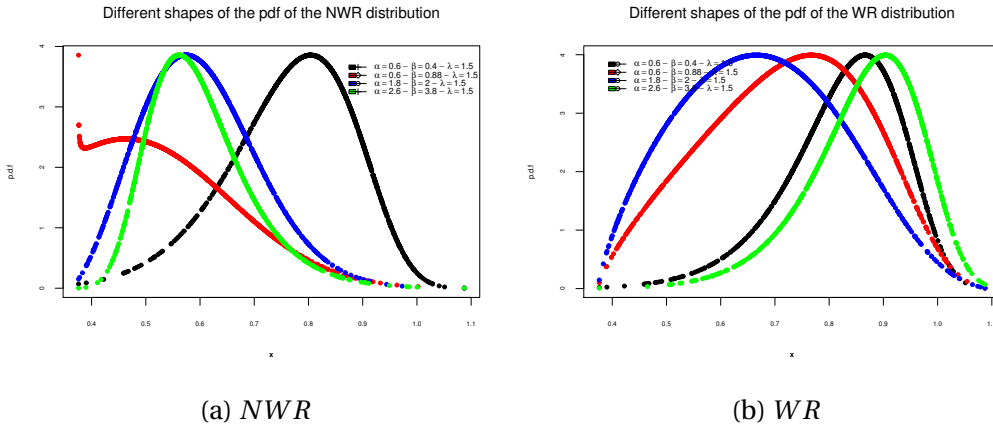


FIGURE 3.1 – pdf du modèle *NWR* vs. *WR*.

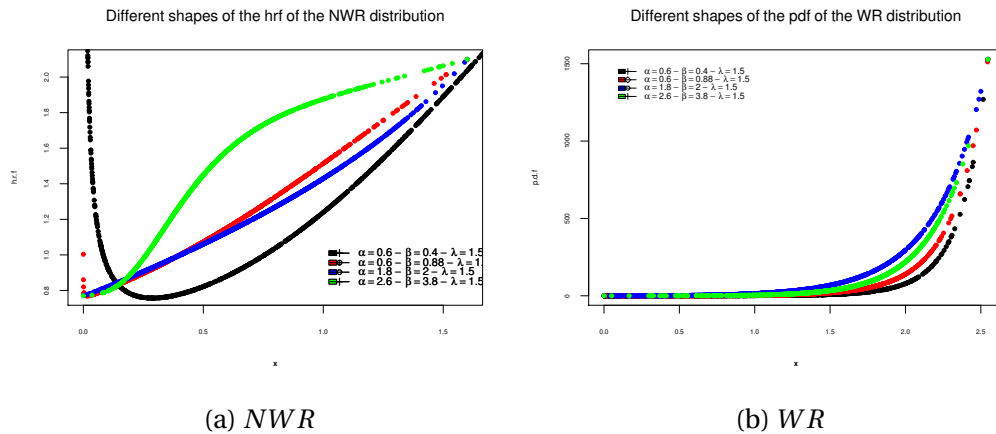


FIGURE 3.2 – hrf du modèle *NWR* vs. *WR*.

Pour différentes combinaisons de paramètres  $\alpha$ ,  $\beta$  et  $\lambda$  de la distribution *NWR*, la Figure 3.1 montre que la distribution *NWR* génère une large gamme de formes de la densité de probabilité, y compris symétrique, asymétrique à gauche et à droite, décroissante et en forme de cloche.

La Figure 3.2 nous permet de voir également la capacité de la famille à produire plusieurs formes du taux de hasard, croissante, baignoire et la forme *S* et *J*.

## 2 Estimation

Dans cette section, nous calculons les estimateurs des paramètres inconnus du modèle *NWR* à l'aide de la méthode du maximum de vraisemblance, dans le cas de données complètes et données censurées à droite.

### 2.1 Estimation des paramètres inconnus dans le cas complet

On suppose  $X_i, (i = \overline{1, n})$ , un  $n$ -échantillon aléatoire provient de la distribution *NWR*, de vecteur de paramètres  $\zeta = (\alpha, \xi, \lambda)$ .

L'équation de la vraisemblance s'écrit :

$$\begin{aligned} L_{NWR}(x_i, \varsigma) &= \prod_{i=1}^n f_{NWR}(x_i, \varsigma) \\ &= \prod_{i=1}^n \left[ 2\alpha\xi \frac{\lambda x_i \exp(-\lambda x_i^2)}{1 - \exp(-\lambda x_i^2)} \{-\log[1 - \exp(-\lambda x_i^2)]\}^{\xi-1} \exp(-\alpha \{-\log[1 - \exp(-\lambda x_i^2)]\}^\xi) \right]. \end{aligned}$$

d'où la log-vraisemblance s'écrit sous la formule suivante :

$$\begin{aligned} l(x_i, \varsigma) &= \sum_{i=1}^n \log f_{NWR}(x_i, \alpha, \xi, \lambda) \\ &= n \log(2\alpha\xi\lambda) + \sum_{i=1}^n \log(x_i) - \sum_{i=1}^n (\lambda x_i^2) + (\xi - 1) \sum_{i=1}^n \log\{-\log[1 - \exp(-\lambda x_i^2)]\} \\ &\quad - \alpha \sum_{i=1}^n \{-\log[1 - \exp(-\lambda x_i^2)]\}^\xi - \sum_{i=1}^n \log[1 - \exp(-\lambda x_i^2)]. \end{aligned}$$

La dérivée de la log-vraisemblance par rapport à chaque paramètre donne un système d'équations appelées les fonctions scores :

$$\begin{aligned} \frac{\partial l_{NWR}(x_i, \varsigma)}{\partial \alpha} &= \frac{n}{\alpha} - \sum_{i=1}^n [-\log(1 - \exp(-\lambda x_i^2))]^\xi, \\ \frac{\partial l_{NWR}(x_i, \varsigma)}{\partial \xi} &= \frac{n}{\xi} + \sum_{i=1}^n \log[-\log(1 - \exp(-\lambda x_i^2))] - \alpha \sum_{i=1}^n \log[K_{NWR}(x_i, \varsigma)] \times K_{NWR}^\xi(x_i, \varsigma), \\ \frac{\partial l_{NWR}(x_i, \varsigma)}{\partial \lambda} &= \frac{n}{\lambda} - \sum_{i=1}^n x_i^2 + (\xi - 1) \sum_{i=1}^n \frac{x_i^2 \exp(-\lambda x_i^2)}{Z_{NWR}(x_i, \varsigma)} \\ &\quad + \alpha \xi \sum_{i=1}^n \frac{x_i^2 \exp(-\lambda x_i^2) K_{NWR}^{\xi-1}(x_i, \varsigma)}{(1 - \exp(-\lambda x_i^2))} - \sum_{i=1}^n \frac{x_i^2 \exp(-\lambda x_i^2)}{(1 - \exp(-\lambda x_i^2))}, \end{aligned}$$

où :

$$\begin{aligned} K_{NWR}(x, \varsigma) &= \{-\log[1 - \exp(-\lambda x^2)]\} \\ Z_{NWR}(x, \varsigma) &= [1 - \exp(-\lambda x^2)] \times \log[1 - \exp(-\lambda x^2)] \end{aligned}$$

Pour résoudre ces équations, on a eu recours à des méthodes numériques en utilisant le logiciel statistique *R*.

## 2.2 La matrice d'information de Fisher estimée dans le cas complet

Les composantes de la matrice d'information de Fisher ainsi obtenues :

$$\begin{aligned}
 \hat{i}_{12} &= -\sum_{i=1}^n \log [K_{NWR}(x_i, \hat{\zeta})] \times [K_{NWR}(x_i, \hat{\zeta})]^{\hat{\xi}}, \\
 \hat{i}_{13} &= \hat{\xi} \sum_{i=1}^n \frac{x_i^2 \exp(-\hat{\lambda}x_i^2) [K_{NWR}(x_i, \hat{\zeta})]^{\hat{\xi}-1}}{(1 - \exp(-\hat{\lambda}x_i^2))}, \\
 \hat{i}_{23} &= \hat{\alpha} \hat{\xi} \sum_{i=1}^n \frac{x_i^2 \exp(-\hat{\lambda}x_i^2) \log [K_{NWR}(x_i, \hat{\zeta})] \times [K_{NWR}(x_i, \hat{\zeta})]^{\hat{\xi}-1}}{1 - \exp(-\hat{\lambda}x_i^2)} \\
 &\quad + \sum_{i=1}^n x_i^2 \exp(-\hat{\lambda}x_i^2) \left\{ \frac{1 - \hat{\alpha} [K_{NWR}(x_i, \hat{\zeta})]^{\hat{\xi}}}{Z_{NWR}(x_i, \hat{\zeta})} \right\}, \\
 \hat{i}_{11} &= -\frac{n}{\hat{\alpha}^2}, \\
 \hat{i}_{22} &= -\frac{n}{\hat{\xi}^2} - \hat{\alpha} \sum_{i=1}^n \log^2 [K_{NWR}(x_i, \hat{\zeta})] \times [K_{NWR}(x_i, \hat{\zeta})]^{\hat{\xi}}, \\
 \hat{i}_{33} &= -\frac{n}{\hat{\lambda}^2} + \sum_{i=1}^n \frac{x_i^4 \exp(-\hat{\lambda}x_i^2)}{1 - \exp(-\hat{\lambda}x_i^2)} \left[ 1 - \hat{\alpha} \hat{\xi} [K_{NWR}(x_i, \hat{\zeta})]^{\hat{\xi}-1} \right] \\
 &\quad + \sum_{i=1}^n \frac{x_i^4 \exp(-2\hat{\lambda}x_i^2)}{[1 - \exp(-\hat{\lambda}x_i^2)]^2} \left\{ 1 - \hat{\alpha} \hat{\xi} K^{\hat{\xi}-1}(x_i, \hat{\zeta}) \left[ 1 + \frac{(\hat{\xi}-1)}{K_{NWR}(x_i, \hat{\zeta})} \right] \right\} \\
 &\quad - (\hat{\xi}-1) \sum_{i=1}^n \frac{x_i^4 \exp(-\hat{\lambda}x_i^2)}{Z_{NWR}(x_i, \hat{\zeta})} \left[ 1 + \frac{\exp(-\hat{\lambda}x_i^2)}{1 - \exp(-\hat{\lambda}x_i^2)} + \frac{\exp(-\hat{\lambda}x_i^2)}{Z_{NWR}(x_i, \hat{\zeta})} \right].
 \end{aligned}$$

## 2.3 Estimation des paramètres inconnus dans le cas censuré

On considère un échantillon censuré à droite constitué de  $n$  observations  $X_i$ ,  $i = \overline{1, n}$ , distribué selon la distribution N-Weibull-Rayleigh, la fonction de vraisemblance du

modèle  $NWR$  s'écrit alors :

$$\begin{aligned} L_{NWR}(x, \varsigma) &= \prod_{i=1}^n h_{NWR}^{\delta_i}(x_i, \varsigma) S_{NWR}(x_i, \varsigma) \\ &= \prod_{i=1}^n \left[ \frac{2\alpha\xi \frac{\lambda x_i \exp(-\lambda x_i^2)}{1 - \exp(-\lambda x_i^2)} \{-\log[1 - \exp(-\lambda x_i^2)]\}^{\xi-1} \times \exp\left(-\alpha \{-\log[1 - \exp(-\lambda x_i^2)]\}^\xi\right)}{1 - \exp\left(-\alpha \{-\log[1 - \exp(-\lambda x_i^2)]\}^\xi\right)} \right]^{\delta_i} \\ &\quad \times \left[ 1 - \exp\left(-\alpha \{-\log[1 - \exp(-\lambda x_i^2)]\}^\xi\right) \right], \end{aligned}$$

alors, la fonction de log-vraisemblance de la distribution  $NWR$  peut être donnée par la formule suivante :

$$\begin{aligned} l_{NWR}(x, \varsigma) &= \sum_{i=1}^n \delta_i [\log(2\alpha\xi\lambda) + \log(x_i) - \lambda x_i^2 + (\xi - 1) \log\{-\log[1 - \exp(-\lambda x_i^2)]\}] \\ &= r \log(2\alpha\xi\lambda) + \sum_{i \in F} \log(x_i) - \lambda \sum_{i \in F} x_i^2 + \sum_{i \in C} \log\left[1 - \exp\left(-\alpha \{-\log[1 - \exp(-\lambda x_i^2)]\}^\xi\right)\right] \\ &\quad - \alpha \sum_{i \in F} \{-\log[1 - \exp(-\lambda x_i^2)]\}^\xi - \sum_{i \in F} \log[1 - \exp(-\lambda x_i^2)] + (\xi - 1) \sum_{i \in F} \log\{-\log[1 - \exp(-\lambda x_i^2)]\}. \end{aligned}$$

Les fonctions scores des paramètres inconnus sont données par :

$$\begin{aligned} \frac{\partial l_{NWR}(x_i, \varsigma)}{\partial \alpha} &= \frac{r}{\alpha} - \sum_{i \in F} K_{NWR}^\xi(x_i, \varsigma) + \sum_{i \in C} \frac{K_{NWR}^\xi(x_i, \varsigma) \exp\left[-\alpha K_{NWR}^\xi(x_i, \varsigma)\right]}{1 - \exp\left[-\alpha K_{NWR}^\xi(x_i, \varsigma)\right]}, \\ \frac{\partial l_{NWR}(x_i, \varsigma)}{\partial \xi} &= \frac{r}{\xi} + \sum_{i \in F} \log[K_{NWR}(x_i, \varsigma)] - \alpha \sum_{i \in F} \log[K_{NWR}(x_i, \varsigma)] K_{NWR}^\xi(x_i, \varsigma) \\ &\quad + \alpha \sum_{i \in C} \frac{\log[K_{NWR}(x_i, \varsigma)] K_{NWR}^\xi(x_i, \varsigma) \exp\left[-\alpha K_{NWR}^\xi(x_i, \varsigma)\right]}{1 - \exp\left[-\alpha K_{NWR}^\xi(x_i, \varsigma)\right]}, \\ \frac{\partial l_{NWR}(x_i, \varsigma)}{\partial \lambda} &= \frac{r}{\lambda} - \sum_{i \in F} x_i^2 - \sum_{i \in F} \frac{x_i^2 \exp(-\lambda x_i^2)}{1 - \exp(-\lambda x_i^2)} + (\xi - 1) \sum_{i \in F} \frac{x_i^2 \exp(-\lambda x_i^2)}{D_{NWR}(x_i, \varsigma)} \\ &\quad + \alpha \xi \frac{x_i^2 K_{NWR}^{\xi-1}(x_i, \varsigma)}{1 - \exp(-\lambda x_i^2)} \left[ \sum_{i \in F} \exp(-\lambda x_i^2) - \sum_{i \in C} \frac{B_{NWR}(x_i, \varsigma)}{1 - \exp\left[-\alpha K_{NWR}^\xi(x_i, \varsigma)\right]} \right]. \end{aligned}$$

Où :

$$\begin{aligned} B_{NWR}(x, \zeta) &= \exp\left(-\alpha \{-\log[1 - \exp(-\lambda x^2)]\}^\xi - \lambda x^2\right), \\ D_{NWR}(x, \zeta) &= [1 - \exp(-\lambda x^2)] \log[1 - \exp(-\lambda x^2)], \\ K_{NWR}(x, \zeta) &= -\log[1 - \exp(-\lambda x^2)]. \end{aligned}$$

En égalant ces fonctions à zéro, on obtient un système d'équations dont les solutions trouvées représentent les estimateurs du maximum de vraisemblance des paramètres inconnus. Cependant, ces équations ne possèdent pas de solutions analytiques, on peut utiliser alors des méthodes itératives.

## 2.4 La matrice d'information de Fisher estimée dans le cas censuré

Les éléments de la matrice d'information de Fisher  $\hat{I} = (\hat{i}_{ll'})$ ,  $l, l' = 1, 2, 3$  sont obtenus comme suit :

$$\begin{aligned} \hat{i}_{11} &= \frac{1}{n} \sum_{i=1}^n \delta_i \left[ \frac{-1}{\hat{\alpha}^2} + \frac{K_{NWR}^{2\hat{\xi}}(x_i, \hat{\zeta}) \exp[-\hat{\alpha} K_{NWR}^{\hat{\xi}}(x_i, \hat{\zeta})]}{\{1 - \exp[-\hat{\alpha} K_{NWR}^{\hat{\xi}}(x_i, \hat{\zeta})]\}^2} \right], \\ \hat{i}_{22} &= \frac{1}{n} \sum_{i=1}^n \delta_i \left[ \frac{-1}{\hat{\xi}^2} - \hat{\alpha} \frac{\log^2 [K_{NWR}(x_i, \hat{\zeta})] K_{NWR}^{\hat{\xi}}(x_i, \hat{\zeta})}{1 - \exp[-\hat{\alpha} K_{NWR}^{\hat{\xi}}(x_i, \hat{\zeta})]} \right] + \\ &\quad \frac{1}{n} \sum_{i=1}^n \delta_i \left[ \hat{\alpha}^2 \frac{\log^2 [K_{NWR}(x_i, \hat{\zeta})] K_{NWR}^{2\hat{\xi}}(x_i, \hat{\zeta}) \exp[-\hat{\alpha} K_{NWR}^{\hat{\xi}}(x_i, \hat{\zeta})]}{\{1 - \exp[-\hat{\alpha} K_{NWR}^{\hat{\xi}}(x_i, \hat{\zeta})]\}^2} \right], \end{aligned}$$

$$\begin{aligned}
 \hat{i}_{33} &= \frac{1}{n} \sum_{i=1}^n \delta_i \left\{ \frac{-1}{\hat{\lambda}^2} + \frac{x_i^4 \exp(-\hat{\lambda}x_i^2)}{1 - \exp(-\hat{\lambda}x_i^2)} \left[ 1 - \frac{(\hat{\xi} - 1)}{\log[1 - \exp(-\hat{\lambda}x_i^2)]} - \hat{\alpha} \hat{\xi} \frac{K_{NWR}^{\hat{\xi}-1}(x_i, \hat{\zeta})}{1 - \exp(-\hat{\lambda}x_i^2)} \right] \right\} \\
 &\quad - \frac{(\hat{\xi} - 1)}{n} \sum_{i=1}^n \delta_i \frac{x_i^4 \exp(-2\hat{\lambda}x_i^2)}{D_{NWR}(x_i, \hat{\zeta})} \left[ \frac{1}{D_{NWR}(x_i, \hat{\zeta})} + \frac{1}{1 - \exp(-\hat{\lambda}x_i^2)} \right] + \\
 &\quad \frac{\hat{\alpha}^2 \hat{\xi}^2}{n} \sum_{i=1}^n \delta_i \frac{K_{NWR}^{2\hat{\xi}-2}(x_i, \hat{\zeta}) h_{NWR}(x_i, \hat{\zeta})}{V_{NWR}(x_i, \hat{\zeta})} \left[ \frac{1}{1 - \exp(-\hat{\lambda}x_i^2)} + \frac{B_{NWR}(x_i, \hat{\zeta})}{V_{NWR}(x_i, \hat{\zeta})} \right] \\
 &\quad - \frac{\hat{\alpha} \hat{\xi}}{n} \sum_{i=1}^n \delta_i \frac{x_i^4 K_{NWR}^{\hat{\xi}-1}(x_i, \hat{\zeta}) B_{NWR}(x_i, \hat{\zeta})}{V_{NWR}(x_i, \hat{\zeta}) [1 - \exp(-\hat{\lambda}x_i^2)]} \left[ 1 + \frac{(\hat{\xi} - 1) \exp(-\hat{\lambda}x_i^2)}{K_{NWR}(x_i, \hat{\zeta})} \right] + \\
 &\quad \frac{1}{n} \sum_{i=1}^n \delta_i \left[ \frac{x_i^4 \exp(-2\hat{\lambda}x_i^2)}{[1 - \exp(-\hat{\lambda}x_i^2)]^2} \left[ 1 - \hat{\alpha} \hat{\xi} (\hat{\xi} - 1) K_{NWR}^{\hat{\xi}-2}(x_i, \hat{\zeta}) \right] \right], \\
 \hat{i}_{12} &= \frac{1}{n} \sum_{i=1}^n \delta_i \frac{\log[K_{NWR}(x_i, \hat{\zeta})] K_{NWR}^{\hat{\xi}}(x_i, \hat{\zeta})}{1 - \exp[-\hat{\alpha} K_{NWR}^{\hat{\xi}}(x_i, \hat{\zeta})]} \left[ \frac{K_{NWR}^{\hat{\xi}}(x_i, \hat{\zeta}) \exp[-\hat{\alpha} K_{NWR}^{\hat{\xi}}(x_i, \hat{\zeta})]}{1 - \exp[-\hat{\alpha} K_{NWR}^{\hat{\xi}}(x_i, \hat{\zeta})]} - 1 \right], \\
 \hat{i}_{13} &= \frac{\hat{\xi}}{n} \sum_{i=1}^n \delta_i \frac{x_i^2 K_{NWR}^{\hat{\xi}-1}(x_i, \hat{\zeta})}{1 - \exp(-\hat{\lambda}x_i^2)} \left[ \exp(-\hat{\lambda}x_i^2) + \frac{B_{NWR}(x_i, \hat{\zeta})}{1 - \exp[-\hat{\alpha} K_{NWR}^{\hat{\xi}}(x_i, \hat{\zeta})]} \right] \\
 &\quad - \frac{\hat{\alpha} \hat{\xi}}{n} \sum_{i=1}^n \delta_i \frac{x_i^2 K_{NWR}^{2\hat{\xi}-1}(x_i, \hat{\zeta}) B_{NWR}(x_i, \hat{\zeta})}{V_{NWR}(x_i, \hat{\zeta})} \left[ \frac{1}{1 - \exp[-\hat{\alpha} K_{NWR}^{\hat{\xi}}(x_i, \hat{\zeta})]} \right], \\
 \hat{i}_{23} &= \frac{1}{n} \sum_{i=1}^n \delta_i \frac{x_i^2 \exp(-\hat{\lambda}x_i^2)}{D_{NWR}(x_i, \hat{\zeta})} - \frac{\hat{\alpha}^2 \hat{\xi}}{n} \sum_{i=1}^n \delta_i \frac{x_i^2 \log[K_{NWR}(x_i, \hat{\zeta})] K_{NWR}^{2\hat{\xi}-1}(x_i, \hat{\zeta}) B_{NWR}(x_i, \hat{\zeta})}{V_{NWR}(x_i, \hat{\zeta}) [1 - \exp[-\hat{\alpha} K_{NWR}^{\hat{\xi}}(x_i, \hat{\zeta})]]} \\
 &\quad + \frac{\hat{\alpha}}{n} \sum_{i=1}^n \delta_i \left\{ x_i^2 K_{NWR}^{\hat{\xi}-1}(x_i, \hat{\zeta}) \left[ \frac{\exp(-\hat{\lambda}x_i^2)}{1 - \exp(-\hat{\lambda}x_i^2)} + \frac{B_{NWR}(x_i, \hat{\zeta})}{V_{NWR}(x_i, \hat{\zeta})} \right] \{1 + \hat{\xi} \log[K_{NWR}(x_i, \hat{\zeta})]\} \right\}.
 \end{aligned}$$

Où :

$$V_{NWR}(x, \zeta) = [1 - \exp(-\lambda x^2)] \times [1 - \exp(-\alpha \{-\log[1 - \exp(-\lambda x^2)]\}^\xi)].$$

## 2.5 Simulation numérique

### Simulation des EMV dans le cas complet

En suivant le même protocole que pour le modèle précédent, nous utilisons le package BB du logiciel *R* pour estimer le vecteur des paramètres  $\zeta$  de la distribution *NWR* et leurs erreurs quadratiques moyennes en effectuant 12,000 simulations pour différentes tailles d'échantillons,  $n = 20, 50, 100, 150, 250, 500, 750$  avec les valeurs initiales ( $\alpha = 2.90, \xi = 3.30, \lambda = 1.80$ ). Les résultats sont illustrés dans la Table 3.1 :

$N = 12,000$	$n = 20$	$n = 50$	$n = 100$	$n = 150$	$n = 250$	$n = 500$	$n = 750$
$\hat{\alpha}$	2.9538	2.9402	2.9333	2.9166	2.9102	2.9055	2.9003
$MSE$	$4.6603 \times 10^{-3}$	$4.0301 \times 10^{-3}$	$3.5478 \times 10^{-3}$	$3.0126 \times 10^{-3}$	$5.3456 \times 10^{-4}$	$3.0211 \times 10^{-4}$	$5.2261 \times 10^{-5}$
$\hat{\xi}$	3.3638	3.3511	3.3409	3.3326	3.3166	3.3018	3.3004
$MSE$	$4.0954 \times 10^{-3}$	$3.2245 \times 10^{-3}$	$3.0645 \times 10^{-3}$	$2.8712 \times 10^{-3}$	$6.7803 \times 10^{-4}$	$4.5870 \times 10^{-4}$	$7.0819 \times 10^{-5}$
$\hat{\lambda}$	1.8538	1.8416	1.8322	1.8266	1.8107	1.8069	1.8001
$MSE$	$7.7508 \times 10^{-3}$	$6.6421 \times 10^{-3}$	$5.8941 \times 10^{-3}$	$4.4587 \times 10^{-3}$	$2.3133 \times 10^{-3}$	$7.0213 \times 10^{-4}$	$1.4503 \times 10^{-4}$

TABLE 3.1 – Les valeurs moyennes des EMV des paramètres et leurs MSE pour les données complètes du modèle *NWR*.

Pour renforcer les résultats de la simulation, on trace dans la Figure 3.3 la courbe de la valeur absolue moyenne des estimateurs du maximum de vraisemblance calculés par rapport aux différentes tailles d'échantillons.

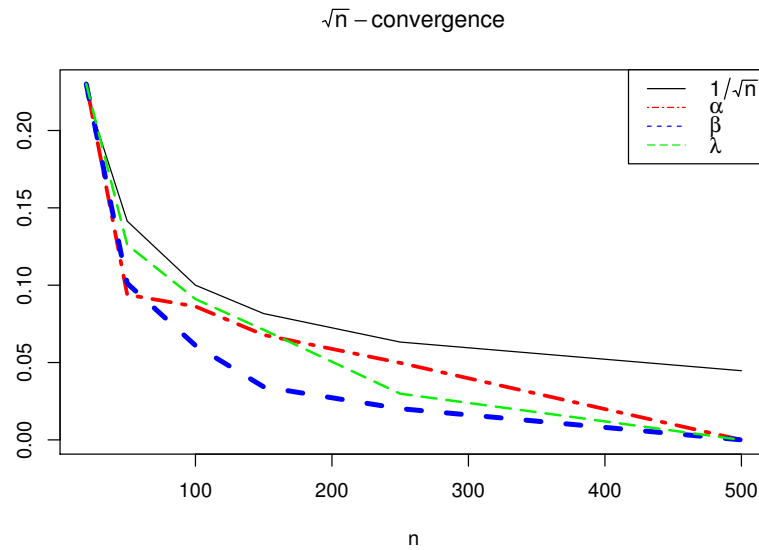


FIGURE 3.3 –  $\sqrt{n}$ -convergence des EMV de la distribution *NWR*.

On peut dire que les estimateurs calculés par la méthode du maximum de vraisemblance de la distribution *NWR* sont  $\sqrt{n}$ -convergeants pour diverses tailles d'échantillon.

### Simulation des EMV censurés

En utilisant le même algorithme (BB), nous simulons 12,000 fois des données censurées à droite de la distribution *NWR*, avec les valeurs initiales des paramètres ( $\alpha = 9,5$ ,  $\xi = 2,5$ ,  $\lambda = 3$ ) pour différentes tailles d'échantillons  $n$ . On calcule les valeurs moyennes des estimateurs du maximum de vraisemblance  $\hat{\alpha}$ ,  $\hat{\xi}$  et  $\hat{\lambda}$  ainsi que les erreurs quadratiques moyennes correspondantes (*MSE*). Les résultats sont résumés dans la Table 3.2 :

$N = 12,000$	$n = 20$	$n = 50$	$n = 100$	$n = 150$	$n = 250$	$n = 500$	$n = 750$
$\hat{\alpha}$	9.6616	9.5756	9.5569	9.5206	9.5173	9.5007	9.4988
$MSE$	$8.7025 \times 10^{-3}$	$9.4710 \times 10^{-4}$	$5.6230 \times 10^{-4}$	$1.4288 \times 10^{-4}$	$6.3754 \times 10^{-5}$	$4.5073 \times 10^{-5}$	$2.1344 \times 10^{-5}$
$\hat{\xi}$	2.6211	2.5988	2.5410	2.5259	2.5102	2.5001	2.4987
$MSE$	$7.4461 \times 10^{-3}$	$2.7906 \times 10^{-3}$	$8.1866 \times 10^{-4}$	$5.8379 \times 10^{-4}$	$1.2121 \times 10^{-4}$	$6.0230 \times 10^{-5}$	$3.0871 \times 10^{-5}$
$\hat{\lambda}$	3.1651	3.0822	3.0529	3.0245	3.0017	2.9984	2.9955
$MSE$	$1.0904 \times 10^{-3}$	$6.0923 \times 10^{-4}$	$2.2929 \times 10^{-4}$	$1.0635 \times 10^{-4}$	$5.4582 \times 10^{-5}$	$3.7117 \times 10^{-5}$	$1.1656 \times 10^{-5}$

TABLE 3.2 – Les valeurs moyennes des EMV des paramètres et leurs MSE pour les données censurées du modèle *NWR*.

Dans la Figure 3.4, la courbe de la valeur absolue moyenne des paramètres estimés par rapport aux différentes tailles d'échantillons  $n$ , montre que les EMV calculés sont  $\sqrt{n}$ -convergentes :

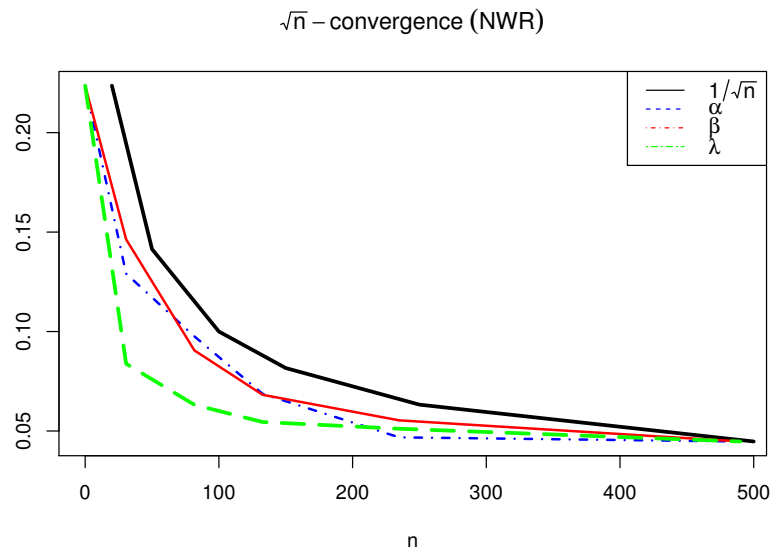


FIGURE 3.4 –  $\sqrt{n}$ -convergence des EMV pour les données censurées de la distribution *NWR*.

### 3 Test d'ajustement pour le modèle *NWR*

#### 3.1 Critère du test d'ajustement dans le cas complet

Nous allons utiliser la statistique développée précédemment par Nikulin-Rao-Robson (1973, 1974) pour construire un critère de test pour le modèle N-Weibull-Rayleigh.

Soit l'hypothèse nulle  $H_0$  selon laquelle le temps de survie  $X_i$  provient de la distribution *NWR* :

$$H_0 : P(X_i \leq x) = F_{NWR}(x, \zeta), \quad \zeta = (\alpha, \xi, \lambda)^T$$

Après avoir calculé les estimateurs du maximum de vraisemblance des paramètres  $\alpha$ ,  $\xi$ ,  $\lambda$  sur un échantillon aléatoire, on regroupe les observations dans  $r$  classes  $I_j$  limitées par  $a_j$ . Ces limites sont choisies telles que les classes obtenues soient équiprobables de probabilité  $p_j$  :

$$p_j(\hat{\zeta}_n) = \int_{a_{j-1}}^{a_j} dF(t, \hat{\zeta}_n) = \frac{1}{r}, \quad j = \overline{1, r}.$$

Alors, les limites estimées  $\hat{a}_j$  des intervalles  $I_j$  sont obtenues comme suit :

$$\hat{a}_j = \left[ \frac{\log \left( 1 - \exp \left\{ - \left[ - \frac{\log \left( \frac{j}{r} \right) ]^{\frac{1}{\xi}}}{\hat{\alpha}} \right\} \right)}{\hat{\lambda}} \right]^{\frac{1}{2}}.$$

Les fréquences empiriques  $\nu = (\nu_1, \dots, \nu_r)^T$  obtenues par ce regroupement sont données comme suit :

$$\nu_j = \text{card} \{i : x_i \in I_j, \} i = \overline{1, n}.$$

### 3.1.1 Les éléments de la statistique du test

Pour calculer la statistique du test  $Y^2$ , on doit d'abord déterminer les éléments qui la composent (voir chapitre 1, section3-1) :

$$\begin{aligned} Y^2 &= X_n^2 + Q, \\ Q &= \frac{1}{n} L^T(\hat{\zeta}) G^{-1} L(\hat{\zeta}). \end{aligned}$$

Les éléments du vecteur  $\hat{L}$  :

$$\begin{aligned} \frac{\partial p_j(\hat{\zeta})}{\partial \hat{\alpha}} &= [K_{NWR}(\hat{a}_j)]^{\hat{\xi}} F_{NWR}(\hat{a}_j) - [K_{NWR}(\hat{a}_{j-1})]^{\hat{\xi}} F_{NWR}(\hat{a}_{j-1}), \\ \frac{\partial p_j(\hat{\zeta})}{\partial \hat{\xi}} &= \hat{\alpha} \log [K_{NWR}(\hat{a}_j)] \times [K_{NWR}(\hat{a}_j)]^{\hat{\xi}} F_{NWR}(\hat{a}_j) - \hat{\alpha} \log [K_{NWR}(\hat{a}_{j-1})] \times [K_{NWR}(\hat{a}_{j-1})]^{\hat{\xi}} F_{NWR}(\hat{a}_{j-1}), \\ \frac{\partial p_j(\hat{\zeta})}{\partial \hat{\lambda}} &= \hat{\alpha} \hat{\xi} \left[ \frac{\hat{a}_{j-1} [K_{NWR}(\hat{a}_{j-1})]^{\hat{\xi}-1} \times B_{NWR}(\hat{a}_{j-1})}{1 - \exp(-\hat{\lambda} \hat{a}_{j-1}^2)} - \frac{\hat{a}_j [K_{NWR}(\hat{a}_j)]^{\hat{\xi}-1} \times B_{NWR}(\hat{a}_j)}{1 - \exp(-\hat{\lambda} \hat{a}_j^2)} \right], j = \overline{1, r} \end{aligned}$$

où :

$$\begin{aligned} K_{NWR}(x, \zeta) &= \exp\left(-\alpha \{-\log[1 - \exp(-\lambda x^2)]\}^{\xi} - \lambda x^2\right), \\ B_{NWR}(x, \zeta) &= \exp\left(-\alpha \{-\log[1 - \exp(-\lambda x^2)]\}^{\xi} - \lambda x^2\right). \end{aligned}$$

Alors, le vecteur  $\hat{\mathbf{L}}_{NWR} = (\hat{L}_1, \dots, \hat{L}_s)^T$  est défini par :

$$\hat{\mathbf{L}}_{NWR}(\hat{\zeta}) = \left[ \hat{L}_1(\hat{\zeta}) = \sum_{j=1}^r \frac{\nu_j}{p_j} \frac{\partial p_j(\hat{\zeta})}{\partial \hat{\alpha}}, \hat{L}_2(\hat{\zeta}) = \sum_{j=1}^r \frac{\nu_j}{p_j} \frac{\partial p_j(\hat{\zeta})}{\partial \hat{\xi}}, \hat{L}_3(\hat{\zeta}) = \sum_{j=1}^r \frac{\nu_j}{p_j} \frac{\partial p_j(\hat{\zeta})}{\partial \hat{\lambda}} \right]^T.$$

Les éléments de la matrice  $\widehat{G}$  sont obtenus comme suit :

$$\begin{aligned}\widehat{g}_{11} &= \widehat{i}_{11} - \sum_{j=1}^r \frac{1}{p_j} \left( \frac{\partial p_j(\widehat{\zeta})}{\partial \widehat{\alpha}} \right)^2, \widehat{g}_{22} = \widehat{i}_{22} - \sum_{j=1}^r \frac{1}{p_j} \left( \frac{\partial p_j(\widehat{\zeta})}{\partial \widehat{\xi}} \right)^2, \\ \widehat{g}_{33} &= \widehat{i}_{33} - \sum_{j=1}^r \frac{1}{p_j} \left( \frac{\partial p_j(\widehat{\zeta})}{\partial \widehat{\lambda}} \right)^2, \widehat{g}_{12} = \widehat{i}_{12} - \sum_{j=1}^r \frac{1}{p_j} \left( \frac{\partial p_j(\widehat{\zeta})}{\partial \widehat{\alpha}} \frac{\partial p_j(\widehat{\zeta})}{\partial \widehat{\xi}} \right), \\ \widehat{g}_{13} &= \widehat{i}_{13} - \sum_{j=1}^r \frac{1}{p_j} \left( \frac{\partial p_j(\widehat{\zeta})}{\partial \widehat{\alpha}} \frac{\partial p_j(\widehat{\zeta})}{\partial \widehat{\lambda}} \right), \widehat{g}_{23} = \widehat{i}_{23} - \sum_{j=1}^r \frac{1}{p_j} \left( \frac{\partial p_j(\widehat{\zeta})}{\partial \widehat{\xi}} \frac{\partial p_j(\widehat{\zeta})}{\partial \widehat{\lambda}} \right), j = \overline{1, r}.\end{aligned}$$

La statistique ainsi obtenue suit une distribution du chi-deux à  $(r - 1)$  degrés de liberté.

### 3.1.2 Simulation des valeurs du critère de test

Afin d'étudier l'efficacité du test d'ajustement pour le modèle *NWR*, on a réalisé une importante étude par simulations numériques. Ainsi et pour tester l'hypothèse nulle  $H_0$  selon laquelle un échantillon de données provient de ce modèle, on génère 12,000 échantillons de tailles différentes ( $n = 20, 50, 100, 150, 250, 750$ ), de la distribution *NWR*. On calcule la valeur moyenne prise par la statistique  $Y^2$ . Ensuite, on compte le pourcentage de rejet de l'hypothèse nulle  $H_0$  par rapport aux risques théoriques correspondants, pour différents seuils de signification  $\epsilon = 1\%, 5\%, 10\%$ , (Table 3.3).

$N = 12,000$

$\epsilon \backslash n$	20	50	100	150	250	500	750
1%	0.0069	0.007	0.009	0.012	0.017	0.011	0.015
5%	0.0340	0.0441	0.0476	0.0491	0.0503	0.0506	0.0510
10%	0.0918	0.0921	0.1019	0.0955	0.1007	0.0901	0.0984

TABLE 3.3 – Comparaison des risques théoriques et des risques empiriques pour les données complètes du modèle *NWR*.

Comme on peut l'observer, les niveaux de signification empiriques de la statistique  $Y^2$  sont semblables aux niveaux théoriques de la distribution du chi-deux à  $(r - 1)$  degrés

de liberté. Par conséquent, nous pouvons dire que le test d'adéquation suggéré est en mesure d'ajuster d'une manière adéquate des ensembles de données de la distribution *NWR*.

### 3.2 Critère du test d'ajustement dans le cas censuré

Pour vérifier l'adéquation du modèle *NWR*, on adapte la statistique du test du type de chi-deux modifiée proposée par Bagdonavicius et Nikulin (2011) à notre modèle. On procède de la même façon que dans le chapitre précédent pour le calcul. Supposons vraie l'hypothèse nulle  $H_0$  qu'un échantillon censuré à droite provienne de la distribution *NWR*. Nous calculons les nombres de défaillances théoriques de manière qu'ils soient tous égaux pour toutes les classes de groupement de données  $I_j$  :

$$I_j = (a_{j-1}, a_j], j = \overline{1, k}, \quad a_0 = 0, \quad a_k = \max(X_{(l)}, \tau).$$

La détermination des bornes des intervalles  $\hat{a}_j$  pour la distribution N-Weibull-Rayleigh s'obtient alors par :

$$\hat{a}_j = \left( -\frac{1}{\hat{\lambda}} \log \left\{ 1 - \exp \left[ - \left( -\frac{1}{\hat{a}} \log \left\{ 1 - \exp \left[ \frac{\sum_{l=1}^{i-1} \Lambda(X_{(l)}, \hat{\zeta}) - E_j}{N - i + 1} \right] \right\} \right]^{\frac{1}{\hat{\zeta}}} \right] \right\} \right)^{\frac{1}{2}}, j = \overline{1, k},$$

où :

$$E_j = (N - i + 1) H(\hat{a}_j, \hat{\zeta}) + \sum_{l=1}^{i-1} H(X_{(l)}, \hat{\zeta}).$$

Le nombre de défaillances observées  $U_j$  et le nombre de défaillances théoriques  $e_j$  sont donnés par :

$$U_j = \sum_{i: X_i \in I_j} \delta_i,$$

$$e_j = \frac{E_k}{k}, \quad E_k = \sum_{i=1}^n H(x_i, \hat{\zeta}).$$

### 3.2.1 Les éléments de la statistique du test

Les éléments de la matrice  $\hat{C} = (\hat{C})_{3 \times k}$  :

$$\hat{C}_{1j} = \frac{1}{n} \sum_{i: X_i \in I_j} \delta_i \left[ \frac{1}{\hat{\alpha}} - \frac{K_{NWR}^{\hat{\xi}}(x_i, \hat{\zeta})}{1 - \exp[-\hat{\alpha} K_{NWR}^{\hat{\xi}}(x_i, \hat{\zeta})]} \right],$$

$$\hat{C}_{2j} = \frac{1}{n} \sum_{i: X_i \in I_j} \delta_i \left[ \frac{1}{\hat{\xi}} + \log[K_{NWR}(x_i, \hat{\zeta})] - \hat{\alpha} \frac{\log[K_{NWR}(x_i, \hat{\zeta})] K_{NWR}^{\hat{\xi}}(x_i, \hat{\zeta})}{1 - \exp[-\hat{\alpha} K_{NWR}^{\hat{\xi}}(x_i, \hat{\zeta})]} \right],$$

$$\hat{C}_{3j} = \frac{1}{n} \sum_{i: X_i \in I_j} \delta_i \left[ \frac{1}{\hat{\lambda}} - x_i^2 + \hat{\alpha} \hat{\xi} \frac{x_i^2 K_{NWR}^{\hat{\xi}-1}(x_i, \hat{\zeta}) B_{NWR}(x_i, \hat{\zeta})}{V_{NWR}(x_i, \hat{\zeta})} \right]$$

$$+ \frac{1}{n} \sum_{i: X_i \in I_j} \delta_i \left[ + \frac{x_i^2 \exp(-\hat{\lambda} x_i^2)}{1 - \exp(-\hat{\lambda} x_i^2)} [\hat{\alpha} \hat{\xi} K_{NWR}^{\hat{\xi}-1}(x_i, \hat{\zeta}) - 1] + (\hat{\xi} - 1) \frac{x_i^2 \exp(-\hat{\lambda} x_i^2)}{D_{NWR}(x_i, \hat{\zeta})} \right].$$

On a :

$$W = (W_1, \dots, W_s)^T = \hat{C} \hat{A}^{-1} Z, \quad \hat{A}_j = \frac{U_j}{n},$$

$$Z_j = \frac{1}{\sqrt{n}} (U_j - e_j), \quad j = \overline{1, k}, \quad l, l' = 1, 2, 3 \text{ et } \hat{\zeta} = (\hat{\alpha}, \hat{\xi}, \hat{\lambda}).$$

Nous obtenons alors la statistique du test  $Y_n^2$  :

$$Y_n^2 = \sum_{i=1}^n \frac{(U_j - e_j)^2}{U_j} + W^T \times \left[ \hat{i}_{ll'} - \sum_{j=1}^k \hat{C}_{lj} \hat{C}_{l'j} \hat{A}_j^{-1} \right]^{-1} \times W.$$

La distribution limite de la statistique obtenue est un chi-deux à  $k$  degrés de liberté.

### 3.2.2 Simulation des valeurs du critère de test

Afin de tester l'adéquation d'une série d'observations au modèle  $NWR$ , à l'aide du logiciel statistique  $R$ , on compare aux risques théoriques ( $\epsilon = 1\%$ ,  $5\%$ ,  $10\%$ ) les risques empiriques du critère de test  $Y_n^2$  pour 12,000 échantillons censurés à droite de différentes tailles,  $n = 20, 50, 100, 150, 250, 500, 750$  (Table 3.4) :

		$N = 12,000$						
$\epsilon \backslash n$		20	50	100	150	250	500	750
$\alpha = 1\%$		0.0076	0.0082	0.0089	0.0091	0.0097	0.0106	0.0124
$\alpha = 5\%$		0.0376	0.0428	0.0487	0.0501	0.0506	0.0512	0.0517
$\alpha = 10\%$		0.0665	0.0722	0.0890	0.0900	0.0935	0.1031	0.1180

TABLE 3.4 – Comparaison des risques théoriques et des risques empiriques pour les données censurées du modèle  $NWR$ .

D'après les résultats obtenus dans la Table 3.4, on peut dire que la statistique  $Y_n^2$  est bien adaptée pour ajuster des données censurées à droite au modèle  $NWR$ .

## 4 Tests classiques

Dans la présente section, nous utilisons le logiciel  $R$  pour calculer les critères d'information classiques suivants  $AIC$ ,  $BIC$ ,  $CAIC$ ,  $HQIC$  et  $KS$  afin d'approuver et valider le

modèle étudié en le comparant avec la distribution Rayleigh ( $R$ ), Inverse Rayleigh ( $IR$ ), Rayleigh-Rayleigh ( $Ra - Ra$ ) et Weibull Rayleigh ( $WR$ ).

Pour cela, nous générons 12,000 échantillons de différentes tailles  $n$  selon toutes ces distributions. On obtient les résultats suivants :

$n$	Model	$AIC$	$BIC$	$CAIC$	$HQIC$	$KS$
10	$R$	39.71	39.96	40.13	39.43	0.760
	$IR$	35.85	36.11	36.27	35.57	0.7624
	$Ra - Ra$	20.70	21.22	22.17	20.13	0.6747
	$WR$	1.13	0.62	3.22	-0.96	0.7591
	<b>NWR</b>	<b>0.80</b>	<b>0.56</b>	<b>3.16</b>	<b>-0.03</b>	<b>0.7588</b>
20	$R$	32.80	32.90	32.82	32.82	0.2521
	$IR$	31.79	31.89	31.819	31.816	0.2522
	$Ra - Ra$	28.51	28.75	28.59	28.56	0.2510
	$WR$	23.09	23.39	23.24	23.15	0.2521
	<b>NWR</b>	<b>22.96</b>	<b>23.26</b>	<b>23.11</b>	<b>23.02</b>	<b>0.2123</b>
30	$R$	113.21	114.61	113.35	113.66	0.1179
	$IR$	100.66	102.06	100.80	101.11	0.1249
	$Ra - Ra$	43.99	46.79	44.43	44.89	0.1186
	$WR$	-36.48	-32.28	-35.56	-35.14	0.1019
	<b>NWR</b>	<b>-31.75</b>	<b>-27.55</b>	<b>-30.83</b>	<b>-30.41</b>	<b>0.1013</b>
50	$R$	187.31	189.22	187.39	188.04	0.0993
	$IR$	165.71	167.63	165.80	166.44	0.0944
	$Ra - Ra$	70.57	74.39	70.82	72.02	0.0952
	$WR$	-65.36	-59.62	-64.83	-63.17	0.0789
	<b>NWR</b>	<b>-56.83</b>	<b>-51.09</b>	<b>-56.31</b>	<b>-54.64</b>	<b>0.0782</b>
100	$R$	375.68	378.29	375.72	376.74	0.0709
	$IR$	331.76	334.37	331.81	332.82	0.0709
	$Ra - Ra$	143.37	148.58	143.50	145.48	0.0739
	$WR$	-127.86	-120.05	-127.61	-124.70	0.0579
	<b>NWR</b>	<b>-113.26</b>	<b>-105.44</b>	<b>-113.01</b>	<b>-110.09</b>	<b>0.0572</b>

TABLE 3.5 – Valeurs des critères d'information  $AIC/BIC/CAIC/HQIC/KS$  pour les données simulées.

Nous remarquons que pour différentes tailles d'échantillons, la distribution  $NWR$  a les

plus petites valeurs des critères de tests utilisés, ce qui nous permet de conclure que le modèle étudié est une bonne alternative aux distributions classiques.

## 5 Applications aux données réelles

Dans cette section, nous allons nous servir de trois jeux de données provenant de domaines différents dans le but de confirmer les résultats théoriques obtenus dans les sections précédentes.

### 5.1 Données sismiques

Supposons que les données sismiques affichées dans Table 3.6, disponible sur "<https://www.donnéesmondiales.com>", sont distribuées selon le modèle *NWR*. Ces observations représentent la profondeur (en kilomètres) causée par le déplacement des plaques tectoniques\* du tremblement de terre massif qui a eu lieu dans l'est de la Turquie et la Syrie en Février 2023 :

10, 17, 16, 10, 15.71, 10, 10, 13.82, 24.79, 7.7, 4.5, 8.4
--

TABLE 3.6 – La profondeur du tremblement de terre en Turquie-Syrie.

\* Un tremblement de terre est le résultat d'une libération brusque d'énergie accumulée par les forces exercées sur les roches, ce qui provoque le déplacement brusque des plaques tectoniques de quelques centimètres à plusieurs kilomètres selon l'intensité de l'énergie libérée.

#### - Analyse graphique :

Nous avons réalisé les graphiques QQ-plot et PP-plot, qui nous permettent respectivement de comparer visuellement les Quantiles théoriques et les Probabilités théoriques calculés à partir de la distribution *NWR* avec la distribution empirique de données.

Nous traçons dans la Figure 3.5, la densité de probabilité estimée, QQ-plot, PP-plot et la fonction de répartition estimée :

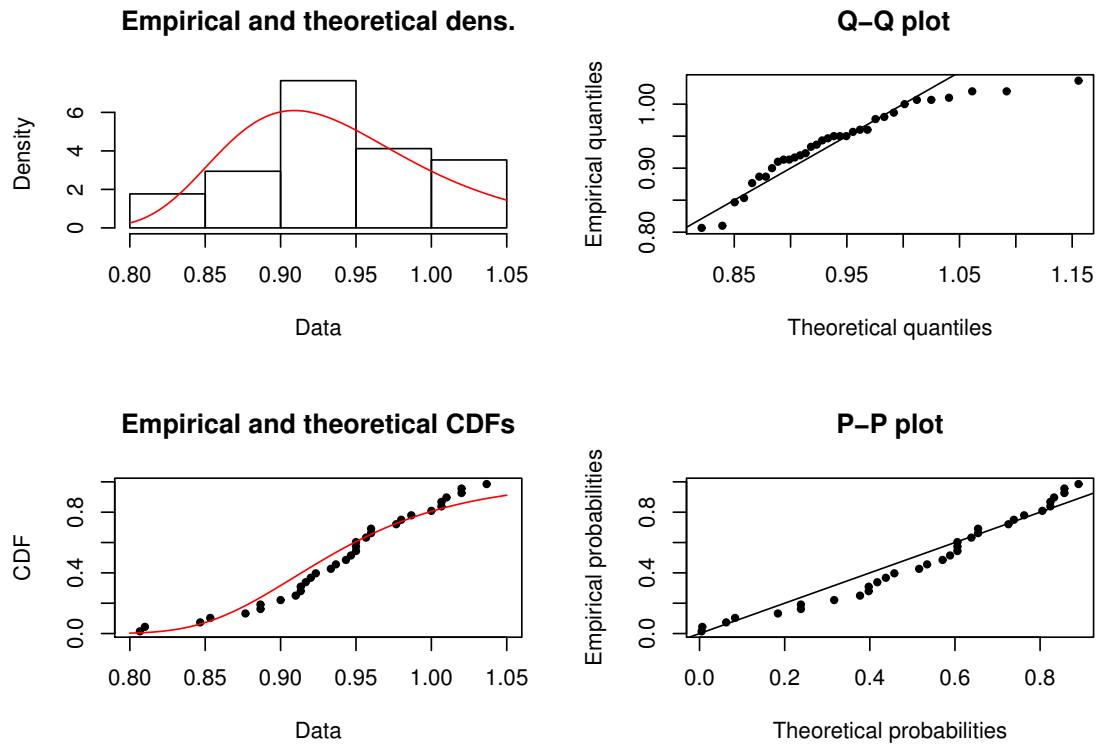


FIGURE 3.5 – pdf, cdf, QQ et PP plot de la profondeur du tremblement de terre en Turquie-Syrie.

Comme nous pouvons le voir (Figure 3.5), la distribution des valeurs des données sismiques utilisées a la même forme que la distribution New-Weibull-Rayleigh que nous avons supposé.

#### - Tests classiques :

Nous utilisons les critères classiques pour confirmer ces résultats. Pour cela, on considère les distributions concurrentes  $WR$ ,  $R$ ,  $IR$ ,  $Ra - Ra$  et  $NWR$ . Les résultats sont

résumés dans la Table 3.7 :

Model	<i>AIC</i>	<i>BIC</i>	<i>CAIC</i>	<i>HQIC</i>	<i>KS</i>
<b>NWR</b>	<b>-9.04</b>	<b>-4.46</b>	<b>-8.24</b>	<b>-7.48</b>	<b>0.1117</b>
<i>R</i>	93.60	95.13	93.73	94.12	0.5212
<i>IR</i>	93.63	95.15	93.75	94.14	0.5548
<i>Ra - Ra</i>	-12.67	9.62	-12.28	-11.63	0.1323
<i>WR</i>	-18.09	-13.51	-17.29	-16.52	0.1166

TABLE 3.7 – Valeurs des critères d'information *AIC*, *BIC*, *CAIC*, *HQIC* et *KS*.

Le modèle *NWR* a les plus petites valeurs des critères d'information classiques, ce qui veut dire que ce modèle est le mieux adapté pour modéliser la profondeur du tremblement de terre qui a frappé la Turquie-Syrie.

**- Le test NRR :**

En plus des tests classiques des modèles de sélection, on calcule la valeur de la statistique  $Y^2$  pour vérifier si ces données sont effectivement distribuées selon *NWR*. À l'aide du logiciel *R*, on trouve les valeurs des *EMV* du vecteur des paramètres  $\zeta$  :

$$\hat{\zeta} = (\hat{\alpha}, \hat{\xi}, \hat{\lambda})^T = (1.32, 4.54, 0.36)^T$$

On regroupe ces données en  $r = 3$  classes. Ensuite, nous calculons les limites estimées des classes  $\hat{a}_j$  et les fréquences empiriques et théoriques  $\hat{v}_j$  et  $p_j$ . Les résultats sont donnés dans la Table 3.8 :

$\hat{a}_j$	0.4017	0.5254	103.1100
$\hat{v}_j$	2	6	4
$p_j$	0.333	0.3333	0.3333

TABLE 3.8 – Valeurs de  $\hat{a}_j$ ,  $\hat{v}_j$  et  $p_j$ .

La valeur de la statistique est  $Y^2 = 2.8819$ . En la comparant à la valeur théorique du chi-deux  $\chi^2$  à  $(3 - 1) = 2$  degrés de liberté pour différents seuils de signification  $\epsilon = 1\%$ ,  $5\%$ ,  $10\%$ , on trouve :

$$Y^2 < \chi_{1\%}^2(3-1) = 9.2103 ; Y^2 < \chi_{5\%}^2(3-1) = 5.9914 ; Y^2 < \chi_{10\%}^2(3-1) = 4.6051.$$

Ceci permet de conclure que la distribution *NWR* peut modéliser la profondeur du tremblement de terre causé par le déplacement des plaques tectoniques à différents seuils de signification  $\epsilon$ .

## 5.2 Données épidémiologiques

Soit l'ensemble de données de Mathers et al., (2013), représentant le taux brut de mortalité (CMR) chez 65 consommateurs de drogues injectables (Table 3.9) :

2.01, 6.32, 3.52, 2.15, 5.42, 2.04, 2.77, 2.26, 1.95, 1.00, 2.45, 0.74, 0.98, 1.27, 2.77, 3.68, 1.18, 1.09, 1.60, 0.57, 3.33, 0.91, 7.14, 2.08, 3.85, 1.99, 7.76, 2.52, 1.57, 4.67, 4.22, 1.92, 1.59, 4.08, 2.02, 0.84, 6.85, 2.18, 2.04, 1.05, 2.91, 1.37, 2.43, 2.28, 3.74, 1.30, 1.59, 1.83, 3.85, 6.30, 4.83, 0.50, 3.40, 2.33, 4.25, 3.49, 2.12, 0.83, 0.54, 3.23, 4.50, 0.71, 0.48, 2.30, 7.73
---

TABLE 3.9 – Le taux de mortalité chez les consommateurs de drogues injectables.

### - Analyse graphique :

La distribution des quantiles et des probabilités théoriques de la distribution *NWR* et la distribution empirique du taux brut de mortalité chez les consommateurs de drogues injectables sont présentées dans la Figure 3.6 (la densité de probabilité, la répartition, le QQ-plot et le PP-plot) :

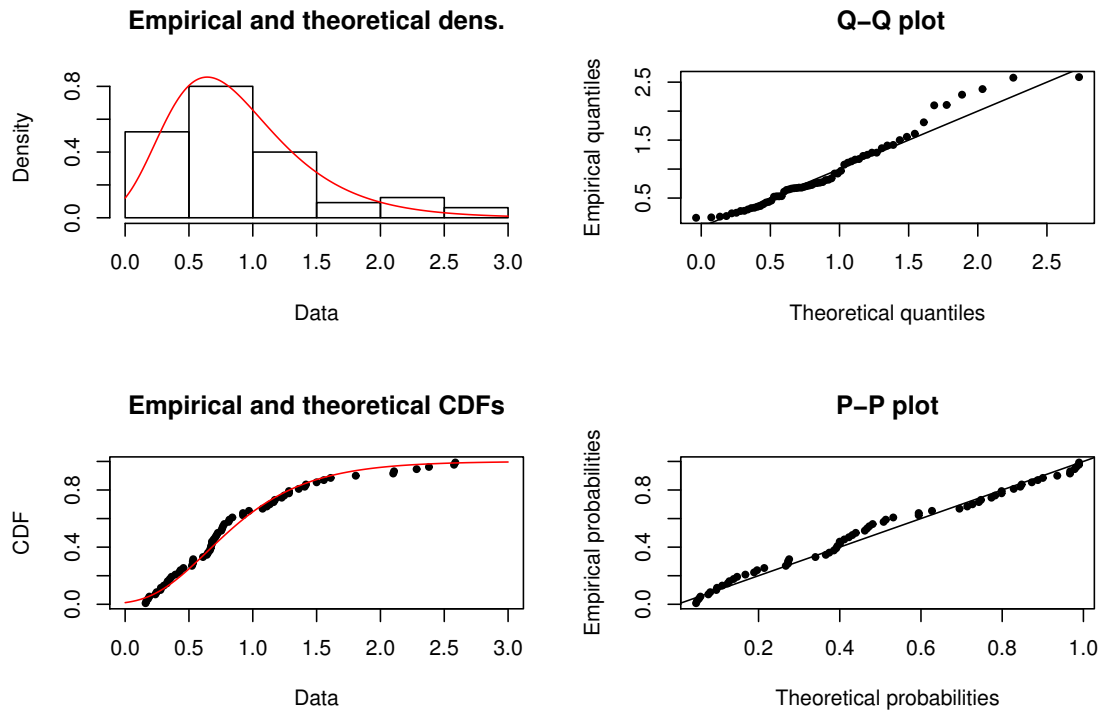


FIGURE 3.6 – pdf, cdf, QQ-plot et PP-plot du taux de mortalité chez les consommateurs de drogues.

Comme le montre la Figure 3.6, la distribution du taux brut de mortalité chez les personnes qui s’injectent des drogues a presque la même forme que les quantiles et les probabilités théoriques de la distribution New-Weibull-Rayleigh.

**- Tests classiques :**

Pour modéliser ces données, la distribution *NWR* a été comparée à cinq distributions alternatives, Rayleigh-Rayleigh (*Ra-Ra*), Gamma Rayleigh (*GaRa*), Marshal Olkin Rayleigh (*MORa*), Truncated-Exponential Skew Symmetric Rayleigh (*TESRa*) et Rayleigh (*Ra*). Pour cela, les différents critères d’information sont calculés sur ces données. Les résultats réunis dans la Table 3.10 montrent que le modèle introduit dans ce chapitre

est très plausible pour décrire le taux de mortalité.

Model	<i>AIC</i>	<i>BIC</i>	<i>CAIC</i>	<i>HQIC</i>	<i>KS</i>
<b>NWR</b>	<b>41.54</b>	<b>48.06</b>	<b>41.93</b>	<b>44.11</b>	<b>0.0896</b>
<i>Ra – Ra</i>	240.98	241.17	245.33	242.70	0.1645
<i>GaRa</i>	244.74	244.93	249.08	246.45	0.2333
<i>MORa</i>	242.09	242.28	246.44	243.81	0.1756
<i>TESRa</i>	243.38	243.58	247.73	245.10	0.1511
<i>Ra</i>	249.30	249.36	251.47	250.16	0.1102

TABLE 3.10 – Valeurs des critères d’information *AIC*, *BIC*, *CAIC*, *HQIC* et *KS*.

**- Le test NRR :**

Nous supposons l’hypothèse nulle  $H_0$  pour laquelle les données épidémiologiques utilisées sont ajustées par la distribution *NWR*. Pour cela, on calcule d’abord les EMV :

$$\hat{\zeta} = (\hat{\alpha}, \hat{\xi}, \hat{\lambda})^T = (0.7594, 3.4829, 0.2950)^T$$

Sous  $H_0$ , on choisit  $r = 8$  intervalles de regroupement  $\hat{I}_j$ , nous calculons donc les limites estimées des classes  $\hat{a}_j$  et les fréquences empiriques  $\hat{v}_j$  et théoriques  $p_j$  correspondantes :

$\hat{a}_j$	1.0597	1.4268	1.8042	2.2472	2.8246	3.6854	5.3442	107.760
$\hat{v}_j$	12	5	4	12	9	7	9	7
$p_j$	0.125	0.125	0.125	0.125	0.125	0.125	0.125	0.125

TABLE 3.11 – Valeurs de  $\hat{a}_j$ ,  $\hat{v}_j$  et  $p_j$ .

Par conséquent, nous obtenons la valeur de la statistique  $Y^2 = 7,4822$  du test NRR et on la compare avec celle de la distribution du chi-deux à 7 degrés de liberté pour différents seuils de signification  $\epsilon = 1\%, 5\%, 10\%$  .

Et on obtient :

$$Y^2 < \chi_{1\%}^2(8-1) = 18.4753 \quad ; \quad Y^2 < \chi_{5\%}^2(7) = 14.0671 \quad ; \quad Y^2 < \chi_{10\%}^2(7) = 12.0170.$$

À partir de ces calculs, on peut conclure que la distribution *NWR* décrit d'une manière adéquate le taux de mortalité chez les consommateurs de drogues injectables.

### 5.3 Données médicales

Considérons l'ensemble de données censurées fournies par Hosmer et Lemeshow (1999) représentant les temps de survie (en mois) de 100 patients qui ont été infectés par le VIH et vérifions si elles peuvent être ajustées par la distribution *NWR*.

5, 6\*, 8, 3, 22, 1\*, 7, 9, 3, 12, 2\*, 12, 1, 15, 34, 1, 4,  
 19\*, 3\*, 2, 2\*, 6, 60\*, 7\*, 60\*, 11, 2\*, 5, 4\*, 1\*, 13, 3\*,  
 2\*, 1\*, 30, 7\*, 4\*, 8\*, 5\*, 10, 2\*, 9\*, 36, 3\*, 9\*, 3\*, 35,  
 8\*, 1\*, 5\*, 11, 56\*, 2\*, 3\*, 15, 1\*, 10, 1\*, 7\*, 3\*, 3\*, 2\*, 32,  
 3\*, 10\*, 11, 3\*, 7\*, 5\*, 31, 5\*, 581\*, 2\*, 1, 3\*, 43, 1\*, 6\*, 53,  
 14, 4\*, 54, 1\*, 1\*, 8\*, 5\*, 1\*, 1\*, 2\*, 7\*, 1\*, 10, 24\*, 7\*, 12\*,  
 4\*, 57, 1\*, 12\*.

TABLE 3.12 – Temps de survie de 100 patients infectés par le VIH.

\* : représente la censure.

#### - Analyse graphique :

Tout d'abord on représente dans la Figure 3.7, les différentes courbes de la densité de probabilité, la répartition, les quantiles et probabilités empiriques de ces observations et théoriques correspondantes au modèle *NWR* :

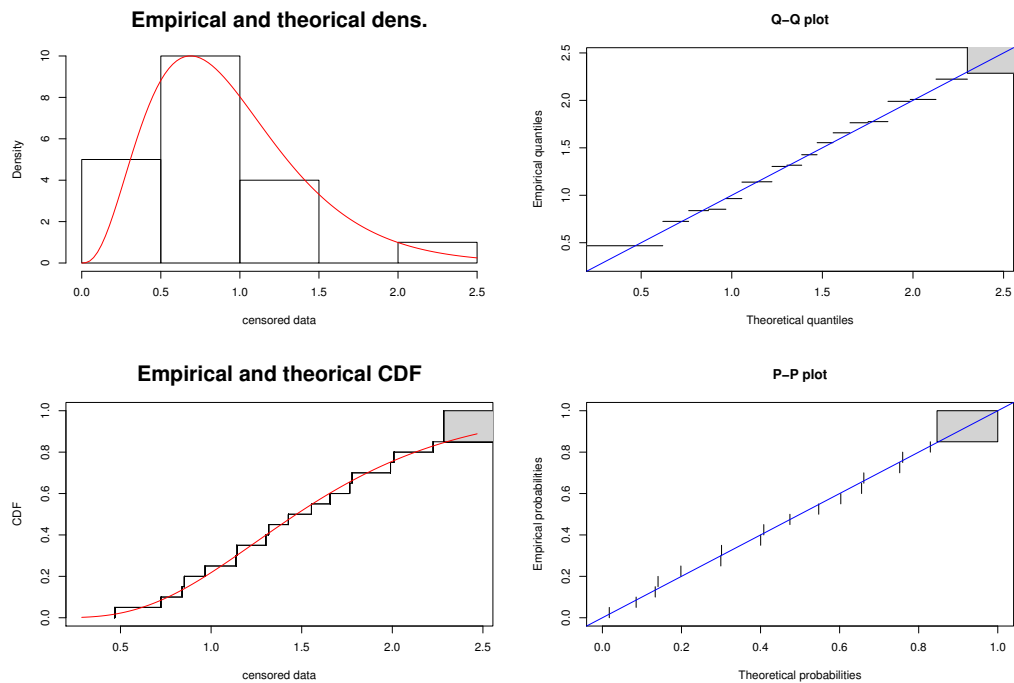


FIGURE 3.7 – pdf, cdf, QQ et PP plot du temps de survie de 100 patients.

Les courbes empiriques des observations coïncident avec celles de la distribution  $NWR$ .

**- Tests classiques :**

Les tests de sélection des modèles classiques  $AIC$ ,  $BIC$ ,  $CAIC$ ,  $HQIC$  et  $KS$  montrent que ce modèle ( $NWR$ ) est une bonne alternative aux distributions  $R$ ,  $IR$ ,  $Ra - Ra$  et  $WR$  (Table 3.13) :

Model	<i>AIC</i>	<i>BIC</i>	<i>CAIC</i>	<i>HQIC</i>	<i>KS</i>
<b>NWR</b>	<b>37.31</b>	<b>38.30</b>	<b>38.81</b>	<b>37.90</b>	<b>0.1095</b>
<i>R</i>	38.84	39.84	39.06	39.04	0.1486
<i>IR</i>	44.56	45.56	44.79	44.76	0.2099
<i>Ra – Ra</i>	38.69	40.68	39.40	39.08	0.1150
<i>WR</i>	39.12	42.11	40.62	39.70	0.1150

TABLE 3.13 – Valeurs des critères d’information *AIC*, *BIC*, *CAIC*, *HQIC* et *KS*.

**- Le test du chi-deux modifié :**

Nous utilisons la statistique de chi-deux modifié construite dans ce travail pour tester ces données censurées à droite au modèle *NWR*. À l’aide du logiciel statistique *R*, nous calculons le vecteur des estimateurs du maximum de vraisemblance  $\hat{\zeta}$  :

$$\hat{\zeta} = (\hat{\alpha}, \hat{\xi}, \hat{\lambda})^T = (2.1308, 0.3515, 1.4651)^T$$

On choisit  $k = 10$  intervalles de regroupement ( $I_j$ ) des données, les calculs intermédiaires de la statistique  $Y_n^2$  du test chi-deux modifié sont présentés dans la Table 3.14 :

$\hat{a}_j$	0.8577	1.3429	1.7532	1.8260	2.0647	2.1931	2.6924	6.2803	4.3664	7.0819
$\hat{U}_j$	4	7	9	16	6	18	5	11	15	9
$e_j$	2.49108	2.49108	2.49108	2.49108	2.49108	2.49108	2.49108	2.49108	2.49108	2.49108
$\hat{C}_{1j}$	-0.0692	0.0595	0.0601	0.1129	-0.01354	1.3102	0.3386	0.1924	-1.2199	0.0417
$\hat{C}_{2j}$	0.6595	1.01080	0.6058	0.929	0.4417	1.1064	0.6279	0.9335	-1.2270	0.2873
$\hat{C}_{3j}$	0.1443	0.2242	0.1405	0.2210	1.0894	0.4306	0.2753	0.3103	1.3061	-0.2756

TABLE 3.14 – Valeurs de  $\hat{a}_j$ ,  $\hat{U}_j$ ,  $e_j$ ,  $\hat{C}_{1j}$ ,  $\hat{C}_{2j}$ ,  $\hat{C}_{3j}$ .

Nous déduisons alors la valeur de  $Y_n^2 = 0.8154$  et la comparons à la distribution du chi-deux  $\chi_\epsilon^2$  pour différents niveaux de signification  $\epsilon = 1\%$ ,  $\epsilon = 5\%$ ,  $\epsilon = 10\%$ , on obtient :

$$Y_n^2 < \chi_{1\%}^2(10) = 21.6659 \quad ; \quad Y_n^2 < \chi_{5\%}^2(10) = 16.9189 \quad ; \quad Y_n^2 < \chi_{10\%}^2(10) = 14.6836.$$

Par conséquent, nous pouvons conclure que la distribution *NWR* modélise de manière fiable le temps de survie des patients infectés par le VIH à un taux de censure assez élevé (62%).

## Modèle à temps de vie accéléré (AFT)

Le modèle à temps de vie accéléré (Accelerated Failure Time) est un type de modèle de régression utilisé dans l'analyse de survie, dans lequel le temps de survie est supposé avoir une relation linéaire avec un ensemble de covariables. Le modèle AFT est un modèle paramétrique qui suppose une distribution spécifique des temps de survie, telles que les distributions Weibull, exponentielle ou la distribution log-normale et la log-logistique, etc.

Le modèle AFT est appelé "accéléré" car il suppose que l'effet des covariables est d'accélérer ou de décélérer le temps de défaillance. En d'autres termes, les covariables affectent la vitesse à laquelle l'événement se produit, mais elles n'affectent pas la forme de la courbe de survie.

Le modèle AFT est largement utilisé dans les domaines médical et l'ingénierie pour étudier le temps de défaillance des appareils ou pour estimer le temps de survie des patients atteints d'une maladie dépendant de certains facteurs. Il est également utilisé dans d'autres domaines, tels que la finance et le marketing pour décrire le temps jusqu'à ce qu'un certain événement se produise (comme la faillite ou l'achat d'un produit).

---

Le principal avantage du modèle AFT est qu'il permet une interprétation aisée des effets des covariables sur la durée de survie.

Du au nombre croissant des données réelles observées lors d'une étude statistique et qui ne peuvent être représentées par les distributions classiques, de nouvelles extensions de distributions existantes sont constamment proposées par les chercheurs dans le domaine de la statistique et de la théorie des probabilités. On cite par exemple le modèle AFT log-gamma introduit par Ortega et al. (2009), le modèle AFT inverse Weibull généralisée étudié par Goual et Seddik-Ameur (2014), le modèle AFT exponentiel exponentié généralisé par Aidi et Seddik-Ameur (2016) et AFT extension de Weibull par Treidi et Seddik-Ameur (2018).

Dans ce chapitre, on propose deux modèles à temps de vie accéléré *AFT – NWW* et *AFT – NWR* dont la distribution de base est respectivement, une *N-Weibull-Weibull* et *N-Weibull-Rayleigh*. Après la présentation des modèles, on utilise la méthode du maximum de vraisemblance pour estimer les paramètres inconnus. Par la suite, on met en place des tests d'adéquation permettant d'ajuster des observations à ce type de modèles de manière satisfaisante.

Considérons que  $E$  est un ensemble de tous les stress ou covariables à  $m$ -dimension possibles et dépendantes du temps.

$$E = \{z(\cdot) = [z_0(\cdot), z_1(\cdot), \dots, z_m(\cdot)]^T\}, \quad z \in [0, \infty[ \in \mathbb{R}^m.$$

On note  $z$  au lieu de  $z(\cdot)$  si la stress est constant et notons par  $E_1 \subset E$  l'ensemble des stress constants. Nous supposons que le temps de défaillance  $X_{z(\cdot)}$  sous le stress  $z(\cdot)$  est une variable aléatoire positive avec fonction de survie :

$$S_{z(\cdot)}(x) = P\{X_{z(\cdot)} > x\}, x > 0, z(\cdot) \in E.$$

Notons par :

$$h_{z(\cdot)}(x) = -\frac{S'_{z(\cdot)}(x)}{S_{z(\cdot)}(x)} \text{ et } H_{z(\cdot)}(x) = -\log S_{z(\cdot)}(x)$$

respectivement, la fonction du taux de hasard et la fonction du taux de hasard cumulé.

Le modèle *AFT* est défini sur l'ensemble  $E$ , si la fonction de survie sous le stress  $z(\cdot) \in E$  vaut :

$$S_{z(\cdot)}(x) = S_0 \left( \int_0^{\tau} r[z(u)] du \right), z(\cdot) \in E.$$

Si  $z(\tau) = z$  est constant, alors le modèle devient :

$$S_{z(\cdot)}(x) = S_0[r(z) t], z \in E_1.$$

La fonction  $r$  peut être choisie parmi une certaine classe de fonction. Souvent, le modèle AFT est paramétré comme :

$$r(z) = e^{-\beta^T z},$$

où  $\beta = (\beta_0, \beta_1, \dots, \beta_m) \in \mathbb{R}^{m+1}$  est le vecteur de paramètres inconnus et

$$z = (z_0, z_1, \dots, z_m)^T$$

le vecteur de stress connu. Alors le modèle AFT paramétrique défini sur l'ensemble  $E$  est donné par la formule :

$$S_{z(\cdot)}(t) = S_0 \left( \int_0^{\tau} \exp[-\beta^T z(u)] du \right), z(\cdot) \in E,$$

et sur l'ensemble des stress constants dans le temps :

$$S(x) = S_0 \left( e^{-\beta^T z} x \right), z \in E.$$

## 1 Construction du modèle AFT-NWW

Nous introduisons un modèle à durée de vie accélérée dont la distribution de base est la distribution N-Weibull-Weibull qu'on notera *AFT – NWW*. La fonction de survie de ce modèle peut être définie comme suit :

$$\begin{aligned} S(x, \vartheta) &= S_0\left(e^{-\beta^T z} x\right) \\ &= 1 - \exp\left[-\alpha\left(-\log\left\{1 - \exp\left[-\lambda\left(xe^{-\beta^T z}\right)^\gamma\right]\right\}\right)^\xi\right], \quad \vartheta = (\alpha, \xi, \lambda, \gamma, \beta_0, \dots, \beta_m) \end{aligned}$$

avec  $\alpha$  est le paramètre d'échelle,  $\xi$  et  $\gamma$  représentent les paramètres de forme et  $\lambda$  le paramètre de position.  $\beta = (\beta_0, \beta_1, \dots, \beta_m)^T$  est le vecteur de paramètres inconnus de la régression des covariables représentant les stress éventuels du modèle *AFT – NWW*

$$z = (z_0, z_1, \dots, z_m)^T$$

Alors, la fonction de répartition, la densité de probabilité, la fonction du taux de hasard et du taux de hasard cumulé sont respectivement exprimées par :

$$\begin{aligned} F_{AFT-NWW}(x, \vartheta) &= \exp\left[-\alpha\left(-\log\left\{1 - \exp\left[-\lambda\left(xe^{-\beta^T z}\right)^\gamma\right]\right\}\right)^\xi\right], \\ f_{AFT-NWW}(x; \vartheta) &= \alpha\xi\lambda\gamma \frac{\left(xe^{-\beta^T z}\right)^{\gamma-1} \exp\left[-\lambda\left(xe^{-\beta^T z}\right)^\gamma\right]}{1 - \exp\left[-\lambda\left(xe^{-\beta^T z}\right)^\gamma\right]} \left(-\log\left\{1 - \exp\left[-\lambda\left(xe^{-\beta^T z}\right)^\gamma\right]\right\}\right)^{(\xi-1)} \\ &\quad \times \exp\left[-\alpha\left(-\log\left\{1 - \exp\left[-\lambda\left(xe^{-\beta^T z}\right)^\gamma\right]\right\}\right)^\xi\right], \\ h_{AFT-NWW}(x; \vartheta) &= \alpha\xi\lambda\gamma \frac{\left(xe^{-\beta^T z}\right)^{\gamma-1} \left(-\log\left\{1 - \exp\left[-\lambda\left(xe^{-\beta^T z}\right)^\gamma\right]\right\}\right)^{(\xi-1)}}{\left\{\exp\left[\lambda\left(xe^{-\beta^T z}\right)^\gamma\right] - 1\right\} \times \left\{\exp\left[\alpha\left(-\log\left\{1 - \exp\left[-\lambda\left(xe^{-\beta^T z}\right)^\gamma\right]\right\}\right)^\xi\right] - 1\right\}}, \\ H_{AFT-NWW}(x; \vartheta) &= -\log\left\{1 - \exp\left[-\alpha\left(-\log\left\{1 - \exp\left[-\lambda\left(xe^{-\beta^T z}\right)^\gamma\right]\right\}\right)^\xi\right]\right\}, \quad \vartheta = (\alpha, \xi, \lambda, \gamma, \beta_0, \beta_1, \dots, \beta_m). \end{aligned}$$

Pour  $\gamma = 1$ , on obtient le cas particulier du modèle *AFT – NWW*, le modèle *AFT N-Weibull-exponentiel (AFT – NWE)*.

### 1.1 Estimation des paramètres inconnus

Pour estimer les paramètres inconnus du modèle *AFT – NWW*, on considère  $X_i = (X_1, \dots, X_n)$  un échantillon aléatoire de distribution *AFT – NWW*. La fonction de log-vraisemblance de  $X_i$  est :

$$\begin{aligned} \ell(x; \vartheta) &= n \log(\alpha \xi \lambda \gamma) + (\gamma - 1) \sum_{i=1}^n \log(x_i e^{-\beta^T z}) - \lambda \sum_{i=1}^n (x_i e^{-\beta^T z})^\gamma - \sum_{i=1}^n \log \left\{ 1 - \exp \left[ -\lambda (x_i e^{-\beta^T z})^\gamma \right] \right\} \\ &\quad - \alpha \sum_{i=1}^n \left( -\log \left\{ 1 - \exp \left[ -\lambda (x_i e^{-\beta^T z})^\gamma \right] \right\} \right)^\xi + (\xi - 1) \sum_{i=1}^n \log \left( -\log \left\{ 1 - \exp \left[ -\lambda (x_i e^{-\beta^T z})^\gamma \right] \right\} \right). \end{aligned}$$

Sans perdre de généralité, nous n'allons considérer que les coefficients  $\beta_0$  et  $\beta_1$  de la régression, le reste se fait de manière analogue. Après de long calculs nous obtenons les fonctions scores pour les paramètres  $\alpha, \xi, \lambda, \gamma, \beta_0$  et  $\beta_1$  :

$$\begin{aligned} \frac{\partial \ell(x_i, \vartheta)}{\partial \alpha} &= \frac{n}{\alpha} - \sum_{i=1}^n [U(x_i, \vartheta)]^\xi, \\ \frac{\partial \ell(x_i, \vartheta)}{\partial \xi} &= \frac{n}{\xi} + \sum_{i=1}^n \log[U(x_i, \vartheta)] \times \left[ 1 - \alpha [U(x_i, \vartheta)]^\xi \right], \\ \frac{\partial \ell(x_i, \vartheta)}{\partial \lambda} &= \frac{n}{\lambda} - \sum_{i=1}^n (x_i e^{-\beta^T z})^\gamma - (\xi - 1) \sum_{i=1}^n \frac{M(x_i, \vartheta)}{D(x_i, \vartheta) \times U(x_i, \vartheta)} - \sum_{i=1}^n \frac{M(x_i, \vartheta)}{D(x_i, \vartheta)} \left[ 1 - \alpha \xi [U(x_i, \vartheta)]^{\xi-1} \right], \\ \frac{\partial \ell(x_i, \vartheta)}{\partial \gamma} &= \frac{n}{\gamma} + \sum_{i=1}^n \log(x_i e^{-\beta^T z}) \left[ 1 - \lambda (x_i e^{-\beta^T z})^\gamma \right] - \lambda (\xi - 1) \sum_{i=1}^n \frac{M(x_i, \vartheta) \log(x_i e^{-\beta^T z})}{D(x_i, \vartheta) \times U(x_i, \vartheta)} \\ &\quad - \lambda \sum_{i=1}^n \frac{M(x_i, \vartheta) \log(x_i e^{-\beta^T z})}{D(x_i, \vartheta)} \left[ 1 - \alpha \xi [U(x_i, \vartheta)]^{\xi-1} \right], \\ \frac{\partial \ell(x_i, \vartheta)}{\partial \beta_0} &= -1 + \lambda (\xi - 1) \gamma \sum_{i=1}^n \frac{M(x_i, \vartheta)}{D(x_i, \vartheta) \times U(x_i, \vartheta)} + \lambda \gamma \sum_{i=1}^n \frac{M(x_i, \vartheta)}{D(x_i, \vartheta)} \left[ 1 - \alpha \xi [U(x_i, \vartheta)]^{\xi-1} \right], \\ \frac{\partial \ell(x_i, \vartheta)}{\partial \beta_1} &= -\sum_{i=1}^n z_i - \lambda (\xi - 1) \gamma \sum_{i=1}^n z_i \frac{M(x_i, \vartheta)}{D(x_i, \vartheta) \times U(x_i, \vartheta)} + \lambda \gamma \sum_{i=1}^n z_i \frac{M(x_i, \vartheta)}{D(x_i, \vartheta)} \left[ 1 - \alpha \xi [U(x_i, \vartheta)]^{\xi-1} \right]. \end{aligned}$$

Où :

$$\begin{aligned} U(x_i, \vartheta) &= -\log\left\{1 - \exp\left[-\lambda\left(xe^{-\beta^T z}\right)^\gamma\right]\right\}, \\ M(x_i, \vartheta) &= \left(xe^{-\beta^T z}\right)^\gamma \exp\left[-\lambda\left(xe^{-\beta^T z}\right)^\gamma\right], \\ D(x_i, \vartheta) &= 1 - \exp\left[-\lambda\left(xe^{-\beta^T z}\right)^\gamma\right]. \end{aligned}$$

Les estimateurs du maximum de vraisemblance du vecteur des paramètres  $\hat{\vartheta} = (\hat{\alpha}, \hat{\xi}, \hat{\lambda}, \hat{\gamma}, \hat{\beta}_0, \hat{\beta}_1)$  sont obtenus en résolvant ce système d'équations non-linéaires par des méthodes numériques.

## 1.2 La matrice d'information de Fisher

Les éléments de la matrice d'information de Fisher  $\hat{I} = (\hat{i})_{6 \times 6}$  sont obtenus comme suit :

$$\begin{aligned} \hat{i}_{11} &= -\frac{n}{\hat{\alpha}^2}, \\ \hat{i}_{22} &= -\frac{n}{\hat{\xi}^2} - \hat{\alpha} \sum_{i=1}^n U(x_i, \hat{\vartheta})^{\hat{\xi}} \log[U(x_i, \hat{\vartheta})], \\ \hat{i}_{33} &= -\frac{n}{\hat{\lambda}^2} + (\hat{\xi} - 1) \sum_{i=1}^n \frac{M(x_i, \hat{\vartheta})^2}{D(x_i, \hat{\vartheta})^2 U(x_i, \hat{\vartheta})^2} \left[1 + U(x_i, \hat{\vartheta}) - \hat{\alpha} \hat{\xi} U(x_i, \hat{\vartheta})^{\hat{\xi}}\right] \\ &\quad + \sum_{i=1}^n \frac{M(x_i, \hat{\vartheta})^2}{D(x_i, \hat{\vartheta})^2} \left[1 - \hat{\alpha} \hat{\xi} U(x_i, \hat{\vartheta})^{\hat{\xi}-1}\right] + \sum_{i=1}^n \frac{M(x_i, \hat{\vartheta})}{D(x_i, \hat{\vartheta}) U(x_i, \hat{\vartheta})} \left[1 + U(x_i, \hat{\vartheta}) - \hat{\alpha} \hat{\xi} U(x_i, \hat{\vartheta})^{\hat{\xi}}\right], \\ \hat{i}_{44} &= -\frac{n}{\hat{\gamma}^2} - \hat{\lambda} \sum_{i=1}^n \left(x_i e^{-\hat{\beta}^T z}\right) \log\left(x_i e^{-\hat{\beta}^T z}\right)^2 + \hat{\lambda}^2 \sum_{i=1}^n \frac{E(x_i, \hat{\vartheta}) M(x_i, \hat{\vartheta})}{D(x_i, \hat{\vartheta})^2} \left[1 - \hat{\alpha} \hat{\xi} U(x_i, \hat{\vartheta})^{\hat{\xi}-1}\right] \\ &\quad + (\hat{\xi} - 1) \hat{\lambda}^2 \sum_{i=1}^n \frac{E(x_i, \hat{\vartheta}) M(x_i, \hat{\vartheta})}{D(x_i, \hat{\vartheta})^2 U(x_i, \hat{\vartheta})^2} \left[1 + U(x_i, \hat{\vartheta}) - \hat{\alpha} \hat{\xi} U(x_i, \hat{\vartheta})^{\hat{\xi}}\right] \\ &\quad - \hat{\lambda} \sum_{i=1}^n \frac{E(x_i, \hat{\vartheta}) M(x_i, \hat{\vartheta})}{D(x_i, \hat{\vartheta})} \left[1 - \hat{\lambda} \left(x_i e^{-\hat{\beta}^T z}\right)^\gamma\right] \times \left[\frac{(\hat{\xi} - 1)}{U(x_i, \hat{\vartheta})} + 1 - \hat{\alpha} \hat{\xi} U(x_i, \hat{\vartheta})^{\hat{\xi}-1}\right], \end{aligned}$$

$$\begin{aligned}
 \hat{i}_{55} &= (\hat{\xi} - 1) \hat{\lambda}^2 \hat{\gamma}^2 \sum_{i=1}^n \frac{M(x_i, \hat{\vartheta})^2}{D(x_i, \hat{\vartheta})^2} \left[ 1 - \hat{\alpha} \hat{\xi} U(x_i, \hat{\vartheta})^{\hat{\xi}} \right] - \hat{\lambda} \hat{\gamma}^2 \sum_{i=1}^n \left( x_i e^{-\hat{\beta}^T z} \right)^{\hat{\gamma}} \\
 &\quad - \hat{\lambda} \hat{\gamma}^2 \sum_{i=1}^n \frac{M(x_i, \hat{\vartheta}) \times \left[ 1 - \hat{\lambda} \left( x_i e^{-\hat{\beta}^T z} \right)^{\hat{\gamma}} \right]}{D(x_i, \hat{\vartheta})} \left[ 1 - \hat{\alpha} \hat{\xi} U(x_i, \hat{\vartheta})^{\hat{\xi}-1} + \frac{(\hat{\xi} - 1)}{U(x_i, \hat{\vartheta})} \right] \\
 &\quad + \hat{\lambda}^2 \hat{\gamma}^2 \sum_{i=1}^n \frac{M(x_i, \hat{\vartheta})^2}{D(x_i, \hat{\vartheta})^2} \left[ 1 - \hat{\alpha} \hat{\xi} U(x_i, \hat{\vartheta})^{\hat{\xi}-1} + \frac{(\hat{\xi} - 1)}{U(x_i, \hat{\vartheta})} \right], \\
 \hat{i}_{66} &= (\hat{\xi} - 1) \hat{\lambda}^2 \hat{\gamma}^2 \sum_{i=1}^n z_i \frac{M(x_i, \hat{\vartheta})^2}{D(x_i, \hat{\vartheta})^2} \left[ 1 - \hat{\alpha} \hat{\xi} U(x_i, \hat{\vartheta})^{\hat{\xi}} \right] - \hat{\lambda} \hat{\gamma}^2 \sum_{i=1}^n z_i \left( x_i e^{-\hat{\beta}^T z} \right)^{\hat{\gamma}} \\
 &\quad - \hat{\lambda} \hat{\gamma}^2 \sum_{i=1}^n z_i \frac{M(x_i, \hat{\vartheta}) \times \left[ 1 - \hat{\lambda} \left( x_i e^{-\hat{\beta}^T z} \right)^{\hat{\gamma}} \right]}{D(x_i, \hat{\vartheta})} \left[ 1 - \hat{\alpha} \hat{\xi} U(x_i, \hat{\vartheta})^{\hat{\xi}-1} + \frac{(\hat{\xi} - 1)}{U(x_i, \hat{\vartheta})} \right] \\
 &\quad + \hat{\lambda}^2 \hat{\gamma}^2 \sum_{i=1}^n z_i \frac{M(x_i, \hat{\vartheta})^2}{D(x_i, \hat{\vartheta})^2} \left[ 1 - \hat{\alpha} \hat{\xi} U(x_i, \hat{\vartheta})^{\hat{\xi}-1} + \frac{(\hat{\xi} - 1)}{U(x_i, \hat{\vartheta})} \right], \\
 \hat{i}_{12} &= - \sum_{i=1}^n U(x_i, \hat{\vartheta})^{\hat{\xi}} \log[U(x_i, \hat{\vartheta})], \\
 \hat{i}_{13} &= \hat{\xi} \sum_{i=1}^n \frac{M(x_i, \hat{\vartheta}) \times U(x_i, \hat{\vartheta})^{\hat{\xi}-1}}{D(x_i, \hat{\vartheta})}, \\
 \hat{i}_{14} &= \hat{\lambda} \hat{\xi} \sum_{i=1}^n \frac{M(x_i, \hat{\vartheta}) \times U(x_i, \hat{\vartheta})^{\hat{\xi}-1}}{D(x_i, \hat{\vartheta})} \log(x_i e^{-\hat{\beta}^T z}), \\
 \hat{i}_{15} &= - \hat{\lambda} \hat{\gamma} \hat{\xi} \sum_{i=1}^n \frac{M(x_i, \hat{\vartheta}) \times U(x_i, \hat{\vartheta})^{\hat{\xi}-1}}{D(x_i, \hat{\vartheta})}, \\
 \hat{i}_{16} &= - \hat{\lambda} \hat{\gamma} \hat{\xi} \sum_{i=1}^n z_i \frac{M(x_i, \hat{\vartheta}) \times U(x_i, \hat{\vartheta})^{\hat{\xi}-1}}{D(x_i, \hat{\vartheta})}, \\
 \hat{i}_{23} &= \hat{\alpha} \sum_{i=1}^n \frac{M(x_i, \hat{\vartheta}) \times U(x_i, \hat{\vartheta})^{\hat{\xi}-1}}{D(x_i, \hat{\vartheta})} \{1 + \hat{\xi} \log[U(x_i, \hat{\vartheta})]\} - \sum_{i=1}^n \frac{M(x_i, \hat{\vartheta})}{D(x_i, \hat{\vartheta}) \times U(x_i, \hat{\vartheta})}, \\
 \hat{i}_{24} &= \hat{\alpha} \hat{\lambda} \sum_{i=1}^n \frac{M(x_i, \hat{\vartheta}) \times U(x_i, \hat{\vartheta})^{\hat{\xi}-1} \log(x_i e^{-\hat{\beta}^T z})}{D(x_i, \hat{\vartheta})} \{1 + \hat{\xi} \log[U(x_i, \hat{\vartheta})]\} \\
 &\quad - \hat{\lambda} \sum_{i=1}^n \frac{M(x_i, \hat{\vartheta}) \log(x_i e^{-\hat{\beta}^T z})}{D(x_i, \hat{\vartheta}) \times U(x_i, \hat{\vartheta})},
 \end{aligned}$$

$$\begin{aligned}
 \hat{i}_{25} &= -\hat{\alpha}\hat{\lambda}\hat{\gamma}\sum_{i=1}^n \frac{M(x_i, \hat{\vartheta}) \times U(x_i, \hat{\vartheta})^{\hat{\xi}-1}}{D(x_i, \hat{\vartheta})} \{1 + \hat{\xi} \log[U(x_i, \hat{\vartheta})]\} + \hat{\lambda}\hat{\gamma}\sum_{i=1}^n \frac{M(x_i, \hat{\vartheta})}{D(x_i, \hat{\vartheta}) \times U(x_i, \hat{\vartheta})}, \\
 \hat{i}_{26} &= -\hat{\alpha}\hat{\lambda}\hat{\gamma}\sum_{i=1}^n z_i \frac{M(x_i, \hat{\vartheta}) \times U(x_i, \hat{\vartheta})^{\hat{\xi}-1}}{D(x_i, \hat{\vartheta})} \{1 + \hat{\xi} \log[U(x_i, \hat{\vartheta})]\} + \hat{\lambda}\hat{\gamma}\sum_{i=1}^n z_i \frac{M(x_i, \hat{\vartheta})}{D(x_i, \hat{\vartheta}) \times U(x_i, \hat{\vartheta})}, \\
 \hat{i}_{34} &= (\hat{\xi}-1)\sum_{i=1}^n \frac{M(x_i, \hat{\vartheta})^2 \log(x_i e^{-\hat{\beta}^T z})}{D(x_i, \hat{\vartheta})^2 U(x_i, \hat{\vartheta})^2} [1 - \hat{\alpha}\hat{\xi}U(x_i, \hat{\vartheta})] - \sum_{i=1}^n (x_i e^{-\hat{\beta}^T z})^{\hat{\gamma}} \log(x_i e^{-\hat{\beta}^T z}) \\
 &\quad - \sum_{i=1}^n \frac{M(x_i, \hat{\vartheta}) \log(x_i e^{-\hat{\beta}^T z})}{D(x_i, \hat{\vartheta})} \left\{ \frac{D(x_i, \hat{\vartheta}) - M(x_i, \hat{\vartheta})}{D(x_i, \hat{\vartheta})} - (x_i e^{-\hat{\beta}^T z})^{\hat{\gamma}} \right\} \times \\
 &\quad \left\{ 1 - \hat{\alpha}\hat{\xi}U(x_i, \hat{\vartheta}) + \frac{\hat{\xi}-1}{U(x_i, \hat{\vartheta})} \right\}, \\
 \hat{i}_{35} &= -\hat{\lambda}(\hat{\xi}-1)\hat{\gamma}\sum_{i=1}^n \frac{M(x_i, \hat{\vartheta})^2}{D(x_i, \hat{\vartheta})^2 U(x_i, \hat{\vartheta})^2} [1 - \hat{\alpha}\hat{\xi}U(x_i, \hat{\vartheta})] + \hat{\gamma}\sum_{i=1}^n (x_i e^{-\hat{\beta}^T z})^{\hat{\gamma}} \\
 &\quad + \hat{\gamma}\sum_{i=1}^n \frac{M(x_i, \hat{\vartheta})}{D(x_i, \hat{\vartheta})} \left\{ 1 + (\hat{\xi}-1)U(x_i, \hat{\vartheta})^{-1} - \hat{\alpha}\hat{\xi}U(x_i, \hat{\vartheta})^{\hat{\xi}-1} \right\} \times \left\{ 1 - \hat{\lambda}(x_i e^{-\hat{\beta}^T z})^{\hat{\gamma}} - \hat{\lambda}\frac{M(x_i, \hat{\vartheta})}{D(x_i, \hat{\vartheta})} \right\}, \\
 \hat{i}_{36} &= -\hat{\lambda}(\hat{\xi}-1)\hat{\gamma}\sum_{i=1}^n z_i \frac{M(x_i, \hat{\vartheta})^2}{D(x_i, \hat{\vartheta})^2 U(x_i, \hat{\vartheta})^2} [1 - \hat{\alpha}\hat{\xi}U(x_i, \hat{\vartheta})] + \hat{\gamma}\sum_{i=1}^n z_i (x_i e^{-\hat{\beta}^T z})^{\hat{\gamma}} \\
 &\quad + \hat{\gamma}\sum_{i=1}^n z_i \frac{M(x_i, \hat{\vartheta})}{D(x_i, \hat{\vartheta})} \left\{ 1 + (\hat{\xi}-1)U(x_i, \hat{\vartheta})^{-1} - \hat{\alpha}\hat{\xi}U(x_i, \hat{\vartheta})^{\hat{\xi}-1} \right\} \left[ 1 - \hat{\lambda}(x_i e^{-\hat{\beta}^T z})^{\hat{\gamma}} - \hat{\lambda}\frac{M(x_i, \hat{\vartheta})}{D(x_i, \hat{\vartheta})} \right], \\
 \hat{i}_{45} &= -1 - \hat{\lambda}^2(\hat{\xi}-1)\hat{\gamma}\sum_{i=1}^n \frac{M(x_i, \hat{\vartheta})^2 \log(x_i e^{-\hat{\beta}^T z})}{D(x_i, \hat{\vartheta})^2 \times U(x_i, \hat{\vartheta})^2} [1 + U(x_i, \hat{\vartheta})] \\
 &\quad - \hat{\lambda}^2\hat{\gamma}\sum_{i=1}^n \frac{M(x_i, \hat{\vartheta})^2 \log(x_i e^{-\hat{\beta}^T z})}{D(x_i, \hat{\vartheta})} \left\{ 1 - \hat{\alpha}\hat{\xi}U(x_i, \hat{\vartheta})^{\hat{\xi}-1} \left[ 1 + \frac{\hat{\xi}-1}{U(x_i, \hat{\vartheta})} \right] \right\} \\
 &\quad + \hat{\lambda}\sum_{i=1}^n \frac{M(x_i, \hat{\vartheta})}{D(x_i, \hat{\vartheta})} \left\{ 1 + \hat{\gamma}\log(x_i e^{-\hat{\beta}^T z}) \times \left[ 1 - \hat{\lambda}(x_i e^{-\hat{\beta}^T z})^{\hat{\gamma}} \right] \right\} \times \left[ 1 + \frac{\hat{\xi}-1}{U(x_i, \hat{\vartheta})} - \hat{\alpha}\hat{\xi}U(x_i, \hat{\vartheta})^{\hat{\xi}-1} \right],
 \end{aligned}$$

$$\begin{aligned}
 \hat{i}_{46} &= -\sum_{i=1}^n z_i - \hat{\lambda}^2 (\hat{\xi} - 1) \hat{\gamma} \sum_{i=1}^n z_i \frac{M(x_i, \hat{\vartheta})^2 \log(x_i e^{-\hat{\beta}^T z})}{D(x_i, \hat{\vartheta})^2 \times U(x_i, \hat{\vartheta})^2} [1 + U(x_i, \hat{\vartheta})] \\
 &\quad - \hat{\lambda}^2 \hat{\gamma} \sum_{i=1}^n z_i \frac{M(x_i, \hat{\vartheta})^2 \log(x_i e^{-\hat{\beta}^T z})}{D(x_i, \hat{\vartheta})} \left\{ 1 - \hat{\alpha} \hat{\xi} U(x_i, \hat{\vartheta})^{\hat{\xi}-1} \left[ 1 + \frac{\hat{\xi}-1}{U(x_i, \hat{\vartheta})} \right] \right\} \\
 &\quad + \hat{\lambda} \sum_{i=1}^n z_i \frac{M(x_i, \hat{\vartheta})}{D(x_i, \hat{\vartheta})} \left\{ 1 + \hat{\gamma} \log(x_i e^{-\hat{\beta}^T z}) \times \left[ 1 - \hat{\lambda} (x_i e^{-\hat{\beta}^T z})^{\hat{\gamma}} \right] \right\} \times \left[ 1 + \frac{\hat{\xi}-1}{U(x_i, \hat{\vartheta})} - \hat{\alpha} \hat{\xi} U(x_i, \hat{\vartheta})^{\hat{\xi}-1} \right], \\
 \hat{i}_{56} &= -\hat{\lambda} \hat{\gamma}^2 \sum_{i=1}^n z_i \frac{M(x_i, \hat{\vartheta}) \left[ \hat{\lambda} (x_i e^{-\hat{\beta}^T z})^{\hat{\gamma}} - 1 \right]}{D(x_i, \hat{\vartheta})} \left\{ 1 + (\hat{\xi}-1) U(x_i, \hat{\vartheta})^{-1} - \hat{\alpha} \hat{\xi} U(x_i, \hat{\vartheta})^{\hat{\xi}-1} \right\} \\
 &\quad - \hat{\alpha} \hat{\xi} \hat{\lambda}^2 \hat{\gamma}^2 \sum_{i=1}^n z_i \frac{M(x_i, \hat{\vartheta})^2 U(x_i, \hat{\vartheta})^{\hat{\xi}-1}}{D(x_i, \hat{\vartheta})^2 U(x_i, \hat{\vartheta})} [U(x_i, \hat{\vartheta}) + \hat{\xi} - 1] - \hat{\lambda} \hat{\gamma}^2 \sum_{i=1}^n z_i (x_i e^{-\hat{\beta}^T z})^{\hat{\gamma}} \\
 &\quad (\hat{\xi}-1) \hat{\lambda}^2 \hat{\gamma} \sum_{i=1}^n z_i \frac{M(x_i, \hat{\vartheta})^2}{D(x_i, \hat{\vartheta})^2 U(x_i, \hat{\vartheta})^2} [1 + U(x_i, \hat{\vartheta})] + \hat{\lambda}^2 \hat{\gamma}^2 \sum_{i=1}^n z_i \frac{M(x_i, \hat{\vartheta})^2}{D(x_i, \hat{\vartheta})^2}.
 \end{aligned}$$

Où :

$$E(x, \vartheta) = \left( x e^{-\beta^T z} \right)^{\gamma} \exp \left[ -\lambda \left( x e^{-\beta^T z} \right)^{\gamma} \right] \log \left( x e^{-\beta^T z} \right)^2, \quad \vartheta = (\alpha, \xi, \lambda, \gamma).$$

### 1.3 Test d'ajustement pour le modèle AFT-NWW

Supposons que les données  $X_1, X_2, \dots, X_n$  sont observées dans un intervalle de temps fini  $[0, \tau]$ . Divisons cet intervalle en  $r$  sous-intervalles  $I_j = ]a_{j-1}, a_j]$ ,  $j = 1, 2, \dots, r$ . Sous l'hypothèse nulle  $H_0$  :

$$H_0 : P(X_i \leq x) = F_{AFT-NWW}(x, \vartheta), \quad \vartheta = (\alpha, \xi, \lambda, \gamma, \beta_0, \beta_1)^T$$

On calcule les estimateurs du maximum de vraisemblance du vecteur des paramètres  $\vartheta$ . Ensuite, nous pouvons calculer les fréquences théoriques telles que :

$$p_j = \int_{a_{j-1}}^{a_j} F(x) dx = \frac{1}{r}, \quad j = 1, \dots, r$$

D'où :

$$\hat{a}_j = e^{-\beta^T z} \left[ \frac{\log \left( 1 - \exp \left\{ - \left[ -\frac{\log\left(\frac{j}{r}\right)}{\hat{\alpha}} \right]^{\frac{1}{\hat{\beta}}} \right\} \right)}{\hat{\lambda}} \right]^{\frac{1}{\hat{\gamma}}}.$$

Soit le vecteur de fréquences empiriques  $v_j$  obtenu en regroupant les observations  $X_1, \dots, X_n$  dans les intervalles  $I_j$  :

$$v_j = \sum_{i=1}^n 1_{\{X_i \in I_j\}}, \quad j = \overline{1, r}.$$

La statistique  $Y^2$  du test  $NRR$  est donnée par :

$$Y^2 = X_n^2 + Q,$$

telles que :

$$X_n^2 = \sum_{j=1}^r \frac{(v_j - np_j)^2}{np_j}, \quad Q = \frac{1}{n} L^T(\hat{\vartheta}) G^{-1} L(\hat{\vartheta}), \quad \hat{\vartheta} = (\hat{\alpha}, \hat{\xi}, \hat{\lambda}, \hat{\gamma}, \hat{\beta}_0, \hat{\beta}_1).$$

#### 1.4 Les éléments de la statistique du test

Pour calculer la statistique  $Y^2$ , nous devons calculer les éléments de la matrice  $\hat{G} = [\hat{g}_{ll'}]_{6 \times 6}$  et le vecteur  $\hat{L}$ .

Les éléments du vecteur  $\hat{L}$  sont obtenus à partir de :

$$p_j(\vartheta) = F_{AFT-NWW}(a_j) - F_{AFT-NWW}(a_{j-1}).$$

Alors :

$$\begin{aligned}
 \frac{\partial p_j}{\partial \hat{\alpha}} &= \sum_{j=1}^r U(\hat{a}_{j-1}, \hat{\vartheta})^{\hat{\xi}} F(\hat{a}_{j-1}, \hat{\vartheta}) - U(\hat{a}_j, \hat{\vartheta})^{\hat{\xi}} F(\hat{a}_j, \hat{\vartheta}), \\
 \frac{\partial p_j}{\partial \hat{\xi}} &= \hat{\alpha} \sum_{j=1}^r U(\hat{a}_{j-1}, \hat{\vartheta})^{\hat{\xi}} \log[U(\hat{a}_{j-1}, \hat{\vartheta})] F(\hat{a}_{j-1}, \hat{\vartheta}) - \hat{\alpha} \sum_{j=1}^r U(\hat{a}_j, \hat{\vartheta})^{\hat{\xi}} \log[U(\hat{a}_j, \hat{\vartheta})] F(\hat{a}_j, \hat{\vartheta}), \\
 \frac{\partial p_j}{\partial \hat{\lambda}} &= \hat{\alpha} \hat{\xi} \sum_{j=1}^r \frac{M(\hat{a}_j, \hat{\vartheta}) U(\hat{a}_j, \hat{\vartheta})^{\hat{\xi}-1} F(\hat{a}_j, \hat{\vartheta})}{D(\hat{a}_j, \hat{\vartheta})} - \hat{\alpha} \hat{\xi} \sum_{j=1}^r \frac{M(\hat{a}_{j-1}, \hat{\vartheta}) U(\hat{a}_{j-1}, \hat{\vartheta})^{\hat{\xi}-1} F(\hat{a}_{j-1}, \hat{\vartheta})}{D(\hat{a}_{j-1}, \hat{\vartheta})}, \\
 \frac{\partial p_j}{\partial \hat{\gamma}} &= \hat{\alpha} \hat{\xi} \hat{\lambda} \sum_{j=1}^r \frac{M(\hat{a}_j, \hat{\vartheta}) U(\hat{a}_j, \hat{\vartheta})^{\hat{\xi}-1} F(\hat{a}_j, \hat{\vartheta})}{D(\hat{a}_j, \hat{\vartheta})} \log(\hat{a}_j e^{-\hat{\beta}^T z}) - \\
 &\quad \hat{\alpha} \hat{\xi} \hat{\lambda} \sum_{j=1}^r \frac{M(\hat{a}_{j-1}, \hat{\vartheta}) U(\hat{a}_{j-1}, \hat{\vartheta})^{\hat{\xi}-1} F(\hat{a}_{j-1}, \hat{\vartheta})}{D(\hat{a}_{j-1}, \hat{\vartheta})} \log(\hat{a}_{j-1} e^{-\hat{\beta}^T z}), \\
 \frac{\partial p_j}{\partial \hat{\beta}_0} &= \hat{\alpha} \hat{\xi} \hat{\lambda} \hat{\gamma} \sum_{j=1}^r \frac{M(\hat{a}_{j-1}, \hat{\vartheta}) U(\hat{a}_{j-1}, \hat{\vartheta})^{\hat{\xi}-1} F(\hat{a}_{j-1}, \hat{\vartheta})}{D(\hat{a}_{j-1}, \hat{\vartheta})} - \hat{\alpha} \hat{\xi} \hat{\lambda} \hat{\gamma} \sum_{j=1}^r \frac{M(\hat{a}_j, \hat{\vartheta}) U(\hat{a}_j, \hat{\vartheta})^{\hat{\xi}-1} F(\hat{a}_j, \hat{\vartheta})}{D(\hat{a}_j, \hat{\vartheta})}, \\
 \frac{\partial p_j}{\partial \hat{\beta}_1} &= \hat{\alpha} \hat{\xi} \hat{\lambda} \hat{\gamma} \sum_{j=1}^r z_i \frac{M(\hat{a}_{j-1}, \hat{\vartheta}) U(\hat{a}_{j-1}, \hat{\vartheta})^{\hat{\xi}-1} F(\hat{a}_{j-1}, \hat{\vartheta})}{D(\hat{a}_{j-1}, \hat{\vartheta})} - \hat{\alpha} \hat{\xi} \hat{\lambda} \hat{\gamma} \sum_{j=1}^r z_i \frac{M(\hat{a}_j, \hat{\vartheta}) U(\hat{a}_j, \hat{\vartheta})^{\hat{\xi}-1} F(\hat{a}_j, \hat{\vartheta})}{D(\hat{a}_j, \hat{\vartheta})}.
 \end{aligned}$$

Où  $F(x, \vartheta)$  est la fonction de répartition du modèle  $AFT-NWW$ . Le vecteur  $\hat{\mathbf{L}}_{AFT-NWW} = (\hat{L}_1, \dots, \hat{L}_6)^T$  est déduit par :

$$\hat{\mathbf{L}}_{AFT-NWW}(\hat{\vartheta}) = \begin{pmatrix} \hat{L}_1 = \sum_{j=1}^r \frac{v_j}{p_j} \frac{\partial p_j(\hat{\vartheta})}{\partial \hat{\alpha}} \\ \hat{L}_2 = \sum_{j=1}^r \frac{v_j}{p_j} \frac{\partial p_j(\hat{\vartheta})}{\partial \hat{\xi}} \\ \hat{L}_3 = \sum_{j=1}^r \frac{v_j}{p_j} \frac{\partial p_j(\hat{\vartheta})}{\partial \hat{\lambda}} \\ \hat{L}_4 = \sum_{j=1}^r \frac{v_j}{p_j} \frac{\partial p_j(\hat{\vartheta})}{\partial \hat{\gamma}} \\ \hat{L}_5 = \sum_{j=1}^r \frac{v_j}{p_j} \frac{\partial p_j(\hat{\vartheta})}{\partial \hat{\beta}_0} \\ \hat{L}_6 = \sum_{j=1}^r \frac{v_j}{p_j} \frac{\partial p_j(\hat{\vartheta})}{\partial \hat{\beta}_1} \end{pmatrix}$$

Les éléments de la matrice  $\widehat{G} = [\widehat{g}_{ll'}]_{6 \times 6}$  sont :

$$\begin{aligned}
 \widehat{g}_{11} &= \widehat{i}_{11} - \sum_{j=1}^r \frac{1}{p_j} \left( \frac{\partial p_j(\widehat{\vartheta})}{\partial \widehat{\alpha}} \right)^2, & \widehat{g}_{22} &= \widehat{i}_{22} - \sum_{j=1}^r \frac{1}{p_j} \left( \frac{\partial p_j(\widehat{\vartheta})}{\partial \widehat{\xi}} \right)^2, \\
 \widehat{g}_{33} &= \widehat{i}_{33} - \sum_{j=1}^r \frac{1}{p_j} \left( \frac{\partial p_j(\widehat{\vartheta})}{\partial \widehat{\lambda}} \right)^2, & \widehat{g}_{44} &= \widehat{i}_{44} - \sum_{j=1}^r \frac{1}{p_j} \left( \frac{\partial p_j(\widehat{\vartheta})}{\partial \widehat{\gamma}} \right)^2, \\
 \widehat{g}_{55} &= \widehat{i}_{55} - \sum_{j=1}^r \frac{1}{p_j} \left( \frac{\partial p_j(\widehat{\vartheta})}{\partial \widehat{\beta}_0} \right)^2, & \widehat{g}_{66} &= \widehat{i}_{66} - \sum_{j=1}^r \frac{1}{p_j} \left( \frac{\partial p_j(\widehat{\vartheta})}{\partial \widehat{\beta}_1} \right)^2, \\
 \widehat{g}_{12} &= \widehat{i}_{12} - \sum_{j=1}^r \frac{1}{p_j} \left( \frac{\partial p_j(\widehat{\vartheta})}{\partial \widehat{\alpha}} \frac{\partial p_j(\widehat{\vartheta})}{\partial \widehat{\xi}} \right), & \widehat{g}_{13} &= \widehat{i}_{13} - \sum_{j=1}^r \frac{1}{p_j} \left( \frac{\partial p_j(\widehat{\vartheta})}{\partial \widehat{\alpha}} \frac{\partial p_j(\widehat{\vartheta})}{\partial \widehat{\lambda}} \right), \\
 \widehat{g}_{14} &= \widehat{i}_{14} - \sum_{j=1}^r \frac{1}{p_j} \left( \frac{\partial p_j(\widehat{\vartheta})}{\partial \widehat{\alpha}} \frac{\partial p_j(\widehat{\vartheta})}{\partial \widehat{\gamma}} \right), & \widehat{g}_{15} &= \widehat{i}_{15} - \sum_{j=1}^r \frac{1}{p_j} \left( \frac{\partial p_j(\widehat{\vartheta})}{\partial \widehat{\alpha}} \frac{\partial p_j(\widehat{\vartheta})}{\partial \widehat{\beta}_0} \right), \\
 \widehat{g}_{16} &= \widehat{i}_{16} - \sum_{j=1}^r \frac{1}{p_j} \left( \frac{\partial p_j(\widehat{\vartheta})}{\partial \widehat{\alpha}} \frac{\partial p_j(\widehat{\vartheta})}{\partial \widehat{\beta}_1} \right), & \widehat{g}_{23} &= \widehat{i}_{23} - \sum_{j=1}^r \frac{1}{p_j} \left( \frac{\partial p_j(\widehat{\vartheta})}{\partial \widehat{\xi}} \frac{\partial p_j(\widehat{\vartheta})}{\partial \widehat{\lambda}} \right), \\
 \widehat{g}_{24} &= \widehat{i}_{24} - \sum_{j=1}^r \frac{1}{p_j} \left( \frac{\partial p_j(\widehat{\vartheta})}{\partial \widehat{\xi}} \frac{\partial p_j(\widehat{\vartheta})}{\partial \widehat{\gamma}} \right), & \widehat{g}_{25} &= \widehat{i}_{25} - \sum_{j=1}^r \frac{1}{p_j} \left( \frac{\partial p_j(\widehat{\vartheta})}{\partial \widehat{\xi}} \frac{\partial p_j(\widehat{\vartheta})}{\partial \widehat{\beta}_0} \right), \\
 \widehat{g}_{26} &= \widehat{i}_{26} - \sum_{j=1}^r \frac{1}{p_j} \left( \frac{\partial p_j(\widehat{\vartheta})}{\partial \widehat{\xi}} \frac{\partial p_j(\widehat{\vartheta})}{\partial \widehat{\beta}_1} \right), & \widehat{g}_{34} &= \widehat{i}_{34} - \sum_{j=1}^r \frac{1}{p_j} \left( \frac{\partial p_j(\widehat{\vartheta})}{\partial \widehat{\lambda}} \frac{\partial p_j(\widehat{\vartheta})}{\partial \widehat{\gamma}} \right), \\
 \widehat{g}_{35} &= \widehat{i}_{35} - \sum_{j=1}^r \frac{1}{p_j} \left( \frac{\partial p_j(\widehat{\vartheta})}{\partial \widehat{\lambda}} \frac{\partial p_j(\widehat{\vartheta})}{\partial \widehat{\beta}_0} \right), & \widehat{g}_{36} &= \widehat{i}_{36} - \sum_{j=1}^r \frac{1}{p_j} \left( \frac{\partial p_j(\widehat{\vartheta})}{\partial \widehat{\lambda}} \frac{\partial p_j(\widehat{\vartheta})}{\partial \widehat{\beta}_1} \right), \\
 \widehat{g}_{45} &= \widehat{i}_{45} - \sum_{j=1}^r \frac{1}{p_j} \left( \frac{\partial p_j(\widehat{\vartheta})}{\partial \widehat{\gamma}} \frac{\partial p_j(\widehat{\vartheta})}{\partial \widehat{\beta}_0} \right), & \widehat{g}_{46} &= \widehat{i}_{46} - \sum_{j=1}^r \frac{1}{p_j} \left( \frac{\partial p_j(\widehat{\vartheta})}{\partial \widehat{\gamma}} \frac{\partial p_j(\widehat{\vartheta})}{\partial \widehat{\beta}_1} \right), \\
 \widehat{g}_{56} &= \widehat{i}_{56} - \sum_{j=1}^r \frac{1}{p_j} \left( \frac{\partial p_j(\widehat{\vartheta})}{\partial \widehat{\beta}_0} \frac{\partial p_j(\widehat{\vartheta})}{\partial \widehat{\beta}_1} \right), & j &= \overline{1, r}.
 \end{aligned}$$

## 1.5 Simulation numérique

### 1.5.1 Les estimateurs du maximum de vraisemblance

Nous générons 12,000 échantillons de différentes tailles ( $n = 15, 50, 100, 150, 250$ ), du modèle  $AFT - NWW$ , avec les valeurs initiales des paramètres ( $\alpha = 1.5, \xi = 0.61, \lambda = 2.4, \gamma = 2.5, \beta_0 = 1.6, \beta_1 = 2$ ). En utilisant l'algorithme de Barzilai-Borwein (BB) du logiciel  $R$ , on calcule les valeurs moyennes des estimateurs  $\widehat{\alpha}, \widehat{\xi}, \widehat{\lambda}, \widehat{\gamma}, \widehat{\beta}_0$  et  $\widehat{\beta}_1$  et leurs

erreurs quadratiques moyennes (MSE) correspondantes (Table 4.1).

Parameters	<b>n = 15</b>	<b>n = 50</b>	<b>n = 100</b>
	<i>AFT - NWW</i>	<i>AFT - NWW</i>	<i>AFT - NWW</i>
$\hat{\alpha}$	1.5124	$\hat{\alpha}$ 1.5084	$\hat{\alpha}$ 1.5001
<i>MSE</i>	$1.8231 \times 10^{-4}$	<i>MSE</i> $1.0332 \times 10^{-4}$	<i>MSE</i> $7.5812 \times 10^{-5}$
$\hat{\xi}$	0.6322	$\hat{\xi}$ 0.6228	$\hat{\xi}$ 0.6173
<i>MSE</i>	$3.5001 \times 10^{-2}$	<i>MSE</i> $3.7413 \times 10^{-3}$	<i>MSE</i> $7.0339 \times 10^{-4}$
$\hat{\lambda}$	2.4182	$\hat{\lambda}$ 2.4101	$\hat{\lambda}$ 2.4071
<i>MSE</i>	$8.7570 \times 10^{-3}$	<i>MSE</i> $1.7579 \times 10^{-3}$	<i>MSE</i> $6.3347 \times 10^{-4}$
$\hat{\gamma}$	2.5699	$\hat{\gamma}$ 2.5304	$\hat{\gamma}$ 2.5203
<i>MSE</i>	$1.9885 \times 10^{-3}$	<i>MSE</i> $1.0053 \times 10^{-3}$	<i>MSE</i> $8.5596 \times 10^{-4}$
$\hat{\beta}_0$	1.6637	$\hat{\beta}_0$ 1.6273	$\hat{\beta}_0$ 1.6198
<i>MSE</i>	$1.7970 \times 10^{-2}$	<i>MSE</i> $5.7994 \times 10^{-3}$	<i>MSE</i> $5.4776 \times 10^{-4}$
$\hat{\beta}_1$	2.2505	$\hat{\beta}_1$ 2.1932	$\hat{\beta}_1$ 2.1337
<i>MSE</i>	$3.1489 \times 10^{-3}$	<i>MSE</i> $8.3304 \times 10^{-4}$	<i>MSE</i> $6.0342 \times 10^{-4}$

Parameters	<b>n = 150</b>	<b>n = 250</b>
	<i>AFT - NWW</i>	<i>AFT - NWW</i>
$\hat{\alpha}$	1.4986	$\hat{\alpha}$ 1.4966
<i>MSE</i>	$5.4344 \times 10^{-5}$	( <i>MSE</i> ) $4.3312 \times 10^{-5}$
$\hat{\xi}$	0.6111	$\hat{\xi}$ 0.6093
<i>MSE</i>	$3.1334 \times 10^{-5}$	<i>MSE</i> $1.1775 \times 10^{-4}$
$\hat{\lambda}$	2.4001	$\hat{\lambda}$ 2.3998
<i>MSE</i>	$1.4706 \times 10^{-5}$	<i>MSE</i> $7.7533 \times 10^{-5}$
$\hat{\gamma}$	2.5123	$\hat{\gamma}$ 2.5088
<i>MSE</i>	$1.3396 \times 10^{-4}$	<i>MSE</i> $5.4498 \times 10^{-5}$
$\hat{\beta}_0$	1.6077	$\hat{\beta}_0$ 1.5913
<i>MSE</i>	$1.6690 \times 10^{-4}$	<i>MSE</i> $7.7966 \times 10^{-5}$
$\hat{\beta}_1$	2.0870	$\hat{\beta}_1$ 2.0061
<i>MSE</i>	$2.7761 \times 10^{-4}$	<i>MSE</i> $3.6617 \times 10^{-5}$

TABLE 4.1 – Les valeurs moyennes des EMV des paramètres et leurs MSE pour le modèle *AFT - NWW*.

Nous observons dans la Table 4.1 que lorsque la taille de l'échantillon  $n$  augmente, les erreurs quadratiques moyennes diminuent et tendent vers zéro. Ce qui nous amène à dire que les estimateurs du maximum de vraisemblance des paramètres obtenus convergent vers les valeurs initiales.

Dans le but de confirmer ces résultats, nous traçons dans la Figure 4.1 les valeurs absolues moyennes des paramètres estimés en fonction de la taille d'échantillon correspondante :

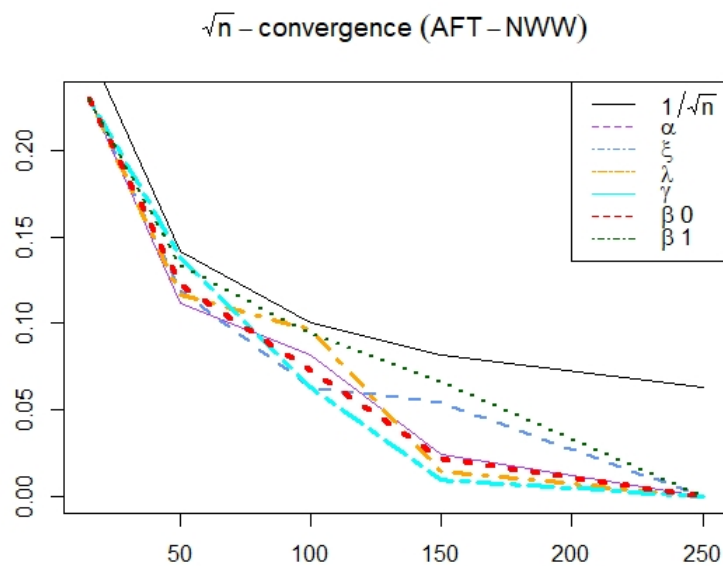


FIGURE 4.1 –  $\sqrt{n}$ -convergence des EMV de la distribution  $AFT - NWW$ .

On remarque que les EMV de la distribution  $AFT - NWW$  sont  $\sqrt{n}$  convergents.

### 1.5.2 La statistique du test

Pour tester l'hypothèse nulle  $H_0$  selon laquelle un échantillon est issu de la distribution  $AFT - NWW$ , nous calculons les valeurs de la statistique  $Y^2$  du test construit dans ce

travail pour 12,000 échantillons simulés de différentes tailles,  $n = 15$ ,  $n = 50$ ,  $n = 100$ ,  $n = 150$  et  $n = 250$ . Pour différents niveaux de signification ( $\epsilon = 1\%$ ,  $5\%$ ,  $10\%$ ), le nombre moyen de cas de rejet de l'hypothèse nulle  $H_0$  est calculé lorsque  $Y^2 > \chi_\epsilon^2(r-1)$ , les résultats sont résumés dans la Table 4.2

<b>n = 15</b>		<b>n = 50</b>		<b>n = 100</b>	
$\epsilon = 1\%$	0.0081	$\epsilon = 1\%$	0.0097	$\epsilon = 1\%$	0.0083
$\epsilon = 5\%$	0.0361	$\epsilon = 5\%$	0.0442	$\epsilon = 5\%$	0.0439
$\epsilon = 10\%$	0.0832	$\epsilon = 10\%$	0.0906	$\epsilon = 10\%$	0.0977

<b>n = 150</b>		<b>n = 250</b>	
$\epsilon = 1\%$	0.0133	$\epsilon = 1\%$	0.0104
$\epsilon = 5\%$	0.0507	$\epsilon = 5\%$	0.0511
$\epsilon = 10\%$	0.1034	$\epsilon = 10\%$	0.1007

TABLE 4.2 – Comparaison des risques théoriques et des risques empiriques du modèle *AFT – NWW*.

Nous observons que les valeurs des niveaux de signification des  $Y^2$  sont proches des valeurs théoriques correspondantes, nous permettant de conclure que le test d'ajustement proposé est bien adapté au modèle *AFT – NWW*.

D'un autre côté, nous représentons dans la Figure 4.2, les histogrammes des statistiques de test calculées  $Y^2$  par rapport à la distribution du chi-deux avec le degré de liberté correspondant.

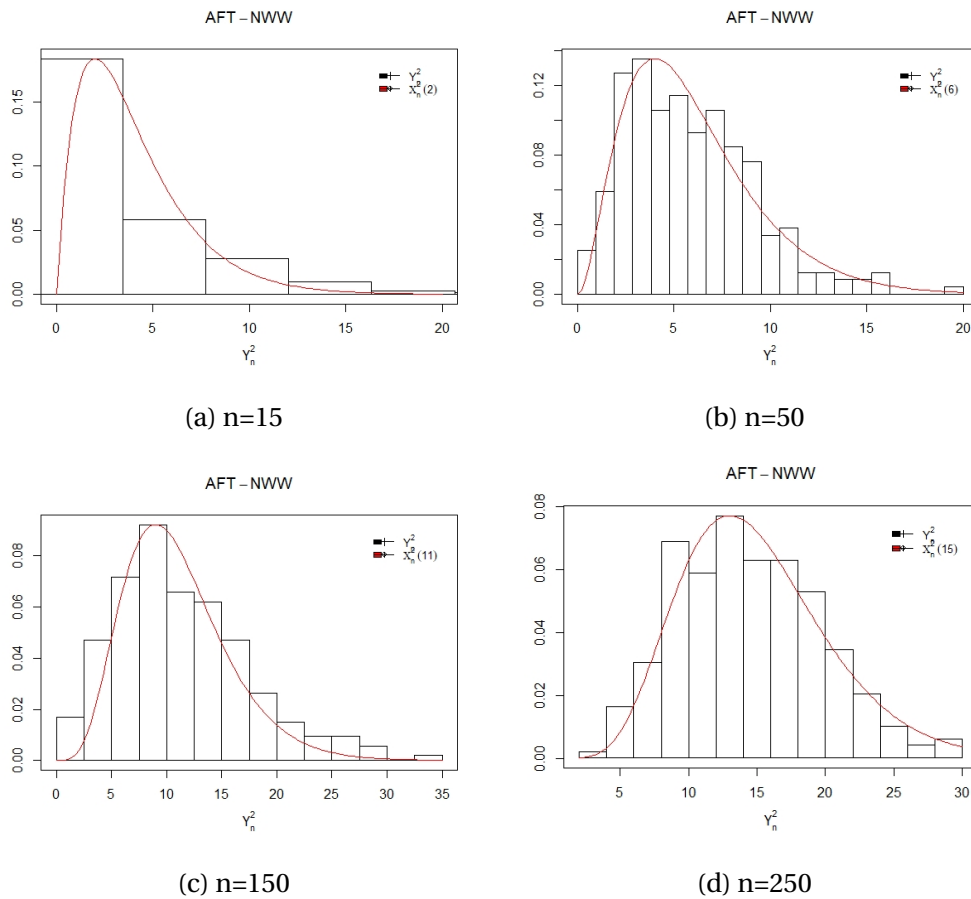


FIGURE 4.2 – Histogrammes de la distribution de  $Y^2$  et la distribution du chi-deux avec le degré de liberté correspondant pour la distribution *AFT - NWW*.

Tenant compte des erreurs des simulations, nous pouvons dire que la distribution limite de la statistique  $Y^2$  est une distribution du chi-deux avec  $(r - 1)$  degrés de liberté.

## 1.6 Application aux données réelles

Considérons l'ensemble de données de *He et al.*, (2012) disponible sur le package 'simex' du logiciel statistique *R* qui donne l'indice de masse corporelle (IMC) de 100 personnes de Busselton en Australie-Occidentale, mesuré à différentes valeurs de la ten-

sion artérielle systolique (Table 4.3). Nous supposons l'hypothèse nulle  $H_0$  pour laquelle ces données peuvent être ajustées par le modèle  $AFT - NWW$ .

$x_i$	24.51, 30.12, 24.53, 28.87, 20.41, 26.42, 21.99, 47.49, 25.21, 21.44...
$z_i$	109, 175, 135, 129, 107, 120, 147, 147, 104, 113...

TABLE 4.3 – L'indice de masse corporelle mesuré à différentes valeurs de la tension artérielle systolique.

Nous utilisons le PP-plot et le QQ-plot pour vérifier l'hypothèse nulle  $H_0$ , (Figure 4.3).

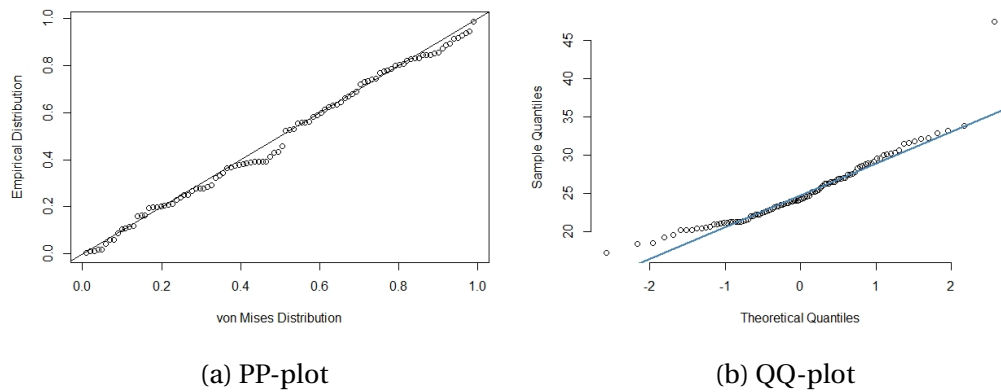


FIGURE 4.3 – PP et QQ plot de l'indice de masse corporelle de 100 personnes.

Comme le montre la Figure 4.3, la distribution des données de l'indice de masse corporelle a presque la même forme que les quantiles et les probabilités théoriques de la répartition du modèle  $AFT - NWW$ .

Sous l'hypothèse nulle  $H_0$ , on calcule le vecteur des EMV du modèle  $AFT - NWW$  :

$$\hat{\theta} = (\hat{\alpha} = 1.4740, \hat{\xi} = 1.9953, \hat{\lambda} = 0.7936, \hat{\gamma} = 0.4239, \hat{\beta}_0 = 0.2267, \hat{\beta}_1 = 0.0383)^T.$$

Si  $r = 10$  intervalles de regroupement  $I_j$  des observations, on calcule les limites  $\hat{a}_j$  de ces intervalles, les fréquences observées  $\hat{v}_j$  et attendues correspondantes  $p_j$ , (Table 4.4) :

$\hat{a}_j$	1.6160	2.9170	4.4715	6.4460	9.0764	1.2780	1.8421	2.8205	5.0576	147.49
$\hat{v}_j$	38	7	13	17	4	9	2	6	3	1
$p_j$	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1

TABLE 4.4 – Valeurs de  $\hat{a}_j$ ,  $\hat{v}_j$  et  $p_j$ .

La matrice d'information de Fisher  $\hat{I}(\hat{\vartheta})$  ainsi obtenue :

$$\hat{I}(\hat{\vartheta}) = \begin{pmatrix} 0.4602 & 0.4938 & -2.6307 & 4.6336 & 0.8851 & 120.3079 \\ 0.4938 & 0.9791 & -2.0718 & 4.5572 & 0.6971 & 105.4088 \\ -2.6307 & -2.0718 & 6.8873 & -4.1526 & 0.1011 & -47.9199 \\ 4.6336 & 4.5572 & -4.1526 & 1.6492 & -0.2904 & -12840.1082 \\ 0.8851 & 0.6971 & 0.1011 & -0.2904 & 0.3853 & -5582.3523 \\ 120.3079 & 105.4088 & -47.9199 & -12840.1082 & -5582.3523 & 51.0477 \end{pmatrix}.$$

Par conséquent, on obtient la valeur du critère de test,  $Y^2 = 8.9740$  et on la compare à la valeur de la statistique du chi-deux  $\chi_{\epsilon}^2$  à 9 degrés de liberté pour différents niveaux de signification  $\epsilon = 1\%$ ,  $\epsilon = 5\%$  et  $\epsilon = 10\%$  :

$$Y^2 < \chi_{0.01}^2(9) = 21.6659,$$

$$Y^2 < \chi_{0.05}^2(9) = 16.9189,$$

$$Y^2 < \chi_{0.1}^2(9) = 14.6836.$$

Les résultats obtenus prouvent que les données utilisées peuvent être modélisées par le modèle *AFT – NWW*. Par conséquent, nous pouvons conclure que la distribution

*AFT – NWW* modélise convenablement des données médicales.

## 2 Construction du modèle AFT-NWR

La distribution *N-Weibull-Rayleigh (NWR)* est une généralisation de la distribution de *Weibull* et celle de *Rayleigh*. L'avantage de cette distribution est qu'elle est polyvalente pour décrire des données positivement asymétriques, unimodales et continues. Cette nouvelle distribution peut être utilisée pour modéliser un large éventail de phénomènes, ceci nous a motivé à introduire un modèle à durée de vie accélérée dont la distribution de base est une *N-Weibull-Rayleigh (AFT – NWR)*. Ce modèle pourra être appliqué dans les tests de vie accélérée en intervenant sur les différents facteurs qui peuvent causer les défaillances.

En suivant le même processus du modèle précédent, nous construisons un modèle à durée de vie accélérée dont la distribution de base est une *N-Weibull-Rayleigh* qu'on note *AFT – NWR*. À l'issue de la présentation du modèle, on utilise la méthode du maximum de vraisemblance pour estimer les paramètres inconnus. Par la suite, en utilisant l'approche proposée par Nikulin-Rao-Robson (1973, 1974) on propose une statistique pour ajuster ce modèle à des observations.

Sous le modèle *AFT – NWR*, la fonction de survie est donnée par la quantité suivante :

$$\begin{aligned} S(x, \zeta) &= S_0\left(e^{-\beta^T z} x\right) \\ &= 1 - \exp\left[-\alpha\left(-\log\left\{1 - \exp\left[-\lambda\left(xe^{-\beta^T z}\right)^2\right]\right\}\right)^\xi\right], \quad \zeta = (\alpha, \xi, \lambda, \beta_0, \beta_1, \dots, \beta_m) \end{aligned}$$

Où  $\alpha$  est le paramètre d'échelle,  $\xi$ ,  $\lambda$  représentent respectivement les paramètres de forme et de position.  $\beta = (\beta_0, \beta_1, \dots, \beta_m)^T$  est le vecteur de paramètres de régression des covariables du modèle *AFT – NWR*.

On définit le modèle *N-Weibull-Rayleigh* par la fonction de répartition, la densité de probabilité et la fonction du taux de hasard et taux de hasard cumulé, données respectivement par :

$$\begin{aligned}
 F_{AFT-NWR}(x; \zeta) &= \exp \left[ -\alpha \left( -\log \left\{ 1 - \exp \left[ -\lambda \left( x e^{-\beta^T z} \right)^2 \right] \right\} \right)^\xi \right], \\
 f_{AFT-NWR}(x; \zeta) &= \alpha \xi \lambda \gamma \frac{\left( x e^{-\beta^T z} \right) \exp \left[ -\lambda \left( x e^{-\beta^T z} \right)^2 \right]}{1 - \exp \left[ -\lambda \left( x e^{-\beta^T z} \right)^2 \right]} \left( -\log \left\{ 1 - \exp \left[ -\lambda \left( x e^{-\beta^T z} \right)^2 \right] \right\} \right)^{(\xi-1)} \\
 &\quad \times \exp \left[ -\alpha \left( -\log \left\{ 1 - \exp \left[ -\lambda \left( x e^{-\beta^T z} \right)^2 \right] \right\} \right)^\xi \right], \\
 h_{AFT-NWR}(x; \zeta) &= 2\alpha \xi \lambda \frac{\left( x e^{-\beta^T z} \right) \left( -\log \left\{ 1 - \exp \left[ -\lambda \left( x e^{-\beta^T z} \right)^2 \right] \right\} \right)^{(\xi-1)}}{\left\{ \exp \left[ \lambda \left( x e^{-\beta^T z} \right)^2 \right] - 1 \right\} \times \left\{ \exp \left[ \alpha \left( -\log \left\{ 1 - \exp \left[ -\lambda \left( x e^{-\beta^T z} \right)^2 \right] \right\} \right)^\xi \right] - 1 \right\}}, \\
 H_{AFT-NWR}(x; \zeta) &= -\log \left[ 1 - \exp \left[ -\alpha \left( -\log \left\{ 1 - \exp \left[ -\lambda \left( x e^{-\beta^T z} \right)^2 \right] \right\} \right)^\xi \right] \right], \quad \vartheta = (\alpha, \xi, \lambda, \beta_0, \beta_1, \dots, \beta_m).
 \end{aligned}$$

## 2.1 Estimation des paramètres inconnus

On considère un échantillon aléatoire  $X = (X_1, \dots, X_n)$  constitué de  $n$  observations provenant de la distribution *AFT-NWR*. L'équation de la log-vraisemblance peut être écrite sous la formule suivante :

$$\begin{aligned}
 \ell(x; \zeta) &= n \log(2\alpha \xi \lambda) + \sum_{i=1}^n \log \left( x_i e^{-\beta^T z} \right) - \lambda \sum_{i=1}^n \left( x_i e^{-\beta^T z} \right)^2 - \sum_{i=1}^n \log \left\{ 1 - \exp \left[ -\lambda \left( x_i e^{-\beta^T z} \right)^2 \right] \right\} \\
 &\quad - \alpha \sum_{i=1}^n \left( -\log \left\{ 1 - \exp \left[ -\lambda \left( x_i e^{-\beta^T z} \right)^2 \right] \right\} \right)^\xi + (\xi - 1) \sum_{i=1}^n \log \left( -\log \left\{ 1 - \exp \left[ -\lambda \left( x_i e^{-\beta^T z} \right)^2 \right] \right\} \right).
 \end{aligned}$$

Sans perdre de généralité, on ne considère que les coefficients  $\beta_0$  et  $\beta_1$  de la régression. La dérivée par rapport à chaque paramètre  $\alpha, \xi, \lambda, \beta_0, \beta_1$  nous permet d'obtenir un

système d'équations non-linéaires :

$$\begin{aligned}\frac{\partial \ell(x_i, \varsigma)}{\partial \alpha} &= \frac{n}{\alpha} - \sum_{i=1}^n [U(x_i, \varsigma)]^\xi, \\ \frac{\partial \ell(x_i, \varsigma)}{\partial \xi} &= \frac{n}{\xi} + \sum_{i=1}^n \log[U(x_i, \varsigma)] \times [1 - \alpha [U(x_i, \varsigma)]^\xi], \\ \frac{\partial \ell(x_i, \varsigma)}{\partial \lambda} &= \frac{n}{\lambda} - \sum_{i=1}^n (x_i e^{-\beta^T z})^2 - (\xi - 1) \sum_{i=1}^n \frac{M(x_i, \varsigma)}{D(x_i, \varsigma) \times U(x_i, \varsigma)} - \sum_{i=1}^n \frac{M(x_i, \varsigma)}{D(x_i, \varsigma)} [1 - \alpha \xi [U(x_i, \varsigma)]^{\xi-1}], \\ \frac{\partial \ell(x_i, \varsigma)}{\partial \beta_0} &= -1 + 2\lambda (\xi - 1) \sum_{i=1}^n \frac{M(x_i, \varsigma)}{D(x_i, \varsigma) \times U(x_i, \varsigma)} + 2\lambda \sum_{i=1}^n \frac{M(x_i, \varsigma)}{D(x_i, \varsigma)} [1 - \alpha \xi [U(x_i, \varsigma)]^{\xi-1}], \\ \frac{\partial \ell(x_i, \varsigma)}{\partial \beta_1} &= -\sum_{i=1}^n z_i + 2\lambda \sum_{i=1}^n z_i \frac{M(x_i, \varsigma)}{D(x_i, \varsigma)} \left\{ \frac{\xi - 1}{U(x_i, \varsigma)} + 1 - \alpha \xi [U(x_i, \varsigma)]^{\xi-1} \right\}.\end{aligned}$$

Où :

$$\begin{aligned}U(x, \varsigma) &= -\log \left\{ 1 - \exp \left[ -\lambda (x e^{-\beta^T z})^2 \right] \right\}, \\ M(x, \varsigma) &= (x e^{-\beta^T z})^2 \exp \left[ -\lambda (x e^{-\beta^T z})^2 \right], \\ D(x, \varsigma) &= 1 - \exp \left[ -\lambda (x e^{-\beta^T z})^2 \right], \quad \varsigma = (\alpha, \xi, \lambda, \beta_0, \beta_1).\end{aligned}$$

La résolution de ce système d'équations non-linéaires nécessite l'utilisation des logiciels de calculs numériques.

## 2.2 La Matrice d'information de Fisher

Les éléments de la matrice d'information de *Fisher* estimée  $\hat{I} = (\hat{i})_{5 \times 5}$  sont :

$$\begin{aligned}\hat{i}_{11} &= -\frac{n}{\hat{\alpha}^2}, \\ \hat{i}_{22} &= -\frac{n}{\hat{\xi}^2} - \hat{\alpha} \sum_{i=1}^n U(x_i, \hat{\varsigma})^{\hat{\xi}} \log[U(x_i, \hat{\varsigma})],\end{aligned}$$

$$\begin{aligned}
 \hat{i}_{33} &= -\frac{n}{\hat{\lambda}^2} + (\hat{\xi} - 1) \sum_{i=1}^n \frac{M(x_i, \hat{\zeta})^2}{D(x_i, \hat{\zeta})^2 U(x_i, \hat{\zeta})^2} \left[ 1 + U(x_i, \hat{\zeta}) - \hat{\alpha} \hat{\xi} U(x_i, \hat{\zeta})^{\hat{\xi}} \right] \\
 &\quad + \sum_{i=1}^n \frac{M(x_i, \hat{\zeta})^2}{D(x_i, \hat{\zeta})^2} \left[ 1 - \hat{\alpha} \hat{\xi} U(x_i, \hat{\zeta})^{\hat{\xi}-1} \right] + \sum_{i=1}^n \frac{M(x_i, \hat{\zeta})}{D(x_i, \hat{\zeta}) U(x_i, \hat{\zeta})} \left[ 1 + U(x_i, \hat{\zeta}) - \hat{\alpha} \hat{\xi} U(x_i, \hat{\zeta})^{\hat{\xi}} \right], \\
 \hat{i}_{44} &= 4(\hat{\xi} - 1) \hat{\lambda}^2 \sum_{i=1}^n \frac{M(x_i, \hat{\zeta})^2}{D(x_i, \hat{\zeta})^2} \left[ 1 - \hat{\alpha} \hat{\xi} U(x_i, \hat{\zeta})^{\hat{\xi}} \right] - 4\hat{\lambda} \sum_{i=1}^n \left( x_i e^{-\hat{\beta}^T z} \right)^2 \\
 &\quad - 4\hat{\lambda} \sum_{i=1}^n \frac{M(x_i, \hat{\zeta}) \times \left[ 1 - \hat{\lambda} \left( x_i e^{-\hat{\beta}^T z} \right)^2 \right]}{D(x_i, \hat{\zeta})} \left[ 1 - \hat{\alpha} \hat{\xi} U(x_i, \hat{\zeta})^{\hat{\xi}-1} + \frac{(\hat{\xi} - 1)}{U(x_i, \hat{\zeta})} \right] \\
 &\quad + 4\hat{\lambda}^2 \sum_{i=1}^n \frac{M(x_i, \hat{\zeta})^2}{D(x_i, \hat{\zeta})^2} \left[ 1 - \hat{\alpha} \hat{\xi} U(x_i, \hat{\zeta})^{\hat{\xi}-1} + \frac{(\hat{\xi} - 1)}{U(x_i, \hat{\zeta})} \right], \\
 \hat{i}_{55} &= 4(\hat{\xi} - 1) \hat{\lambda}^2 \sum_{i=1}^n z_i \frac{M(x_i, \hat{\zeta})^2}{D(x_i, \hat{\zeta})^2} \left[ 1 - \hat{\alpha} \hat{\xi} U(x_i, \hat{\zeta})^{\hat{\xi}} \right] - 4\hat{\lambda} \sum_{i=1}^n z_i \left( x_i e^{-\hat{\beta}^T z} \right)^2 \\
 &\quad - 4\hat{\lambda} \sum_{i=1}^n z_i \frac{M(x_i, \hat{\zeta}) \times \left[ 1 - \hat{\lambda} \left( x_i e^{-\hat{\beta}^T z} \right)^2 \right]}{D(x_i, \hat{\zeta})} \left[ 1 - \hat{\alpha} \hat{\xi} U(x_i, \hat{\zeta})^{\hat{\xi}-1} + \frac{(\hat{\xi} - 1)}{U(x_i, \hat{\zeta})} \right] \\
 &\quad + 4\hat{\lambda}^2 \sum_{i=1}^n z_i \frac{M(x_i, \hat{\zeta})^2}{D(x_i, \hat{\zeta})^2} \left[ 1 - \hat{\alpha} \hat{\xi} U(x_i, \hat{\zeta})^{\hat{\xi}-1} + \frac{(\hat{\xi} - 1)}{U(x_i, \hat{\zeta})} \right], \\
 \hat{i}_{12} &= -\sum_{i=1}^n U(x_i, \hat{\zeta})^{\hat{\xi}} \log[U(x_i, \hat{\zeta})], \\
 \hat{i}_{13} &= \hat{\xi} \sum_{i=1}^n \frac{M(x_i, \hat{\zeta}) \times U(x_i, \hat{\zeta})^{\hat{\xi}-1}}{D(x_i, \hat{\zeta})}, \\
 \hat{i}_{14} &= -2\hat{\lambda} \hat{\xi} \sum_{i=1}^n \frac{M(x_i, \hat{\zeta}) \times U(x_i, \hat{\zeta})^{\hat{\xi}-1}}{D(x_i, \hat{\zeta})}, \\
 \hat{i}_{15} &= -2\hat{\lambda} \hat{\xi} \sum_{i=1}^n z_i \frac{M(x_i, \hat{\zeta}) \times U(x_i, \hat{\zeta})^{\hat{\xi}-1}}{D(x_i, \hat{\zeta})}, \\
 \hat{i}_{23} &= \hat{\alpha} \sum_{i=1}^n \frac{M(x_i, \hat{\zeta}) \times U(x_i, \hat{\zeta})^{\hat{\xi}-1}}{D(x_i, \hat{\zeta})} \{1 + \hat{\xi} \log[U(x_i, \hat{\zeta})]\} - \sum_{i=1}^n \frac{M(x_i, \hat{\zeta})}{D(x_i, \hat{\zeta}) \times U(x_i, \hat{\zeta})}, \\
 \hat{i}_{24} &= -2\hat{\alpha} \hat{\lambda} \sum_{i=1}^n \frac{M(x_i, \hat{\zeta}) \times U(x_i, \hat{\zeta})^{\hat{\xi}-1}}{D(x_i, \hat{\zeta})} \{1 + \hat{\xi} \log[U(x_i, \hat{\zeta})]\} + 2\hat{\lambda} \sum_{i=1}^n \frac{M(x_i, \hat{\zeta})}{D(x_i, \hat{\zeta}) \times U(x_i, \hat{\zeta})}, \\
 \hat{i}_{25} &= -2\hat{\alpha} \hat{\lambda} \sum_{i=1}^n z_i \frac{M(x_i, \hat{\zeta}) \times U(x_i, \hat{\zeta})^{\hat{\xi}-1}}{D(x_i, \hat{\zeta})} \{1 + \hat{\xi} \log[U(x_i, \hat{\zeta})]\} + 2\hat{\lambda} \sum_{i=1}^n z_i \frac{M(x_i, \hat{\zeta})}{D(x_i, \hat{\zeta}) \times U(x_i, \hat{\zeta})},
 \end{aligned}$$

$$\begin{aligned}
 \hat{i}_{34} &= -2\hat{\lambda}(\hat{\xi}-1) \sum_{i=1}^n \frac{M(x_i, \hat{\zeta})^2}{D(x_i, \hat{\zeta})^2 U(x_i, \hat{\zeta})^2} [1 - \hat{\alpha}\hat{\xi}U(x_i, \hat{\zeta})] + 2 \sum_{i=1}^n (x_i e^{-\hat{\beta}^T z})^2 \\
 &\quad + 2 \sum_{i=1}^n \frac{M(x_i, \hat{\zeta})}{D(x_i, \hat{\zeta})} \left[ 1 + (\hat{\xi}-1)U(x_i, \hat{\zeta})^{-1} - \hat{\alpha}\hat{\xi}U(x_i, \hat{\zeta})^{\hat{\xi}-1} \right] \left[ 1 - \hat{\lambda}(x_i e^{-\hat{\beta}^T z})^2 - \hat{\lambda} \frac{M(x_i, \hat{\zeta})}{D(x_i, \hat{\zeta})} \right], \\
 \hat{i}_{35} &= -2\hat{\lambda}(\hat{\xi}-1) \sum_{i=1}^n z_i \frac{M(x_i, \hat{\zeta})^2}{D(x_i, \hat{\zeta})^2 U(x_i, \hat{\zeta})^2} [1 - \hat{\alpha}\hat{\xi}U(x_i, \hat{\zeta})] + 2 \sum_{i=1}^n z_i (x_i e^{-\hat{\beta}^T z})^2 \\
 &\quad + 2 \sum_{i=1}^n z_i \frac{M(x_i, \hat{\zeta})}{D(x_i, \hat{\zeta})} \left[ 1 + (\hat{\xi}-1)U(x_i, \hat{\zeta})^{-1} - \hat{\alpha}\hat{\xi}U(x_i, \hat{\zeta})^{\hat{\xi}-1} \right] \left[ 1 - \hat{\lambda}(x_i e^{-\hat{\beta}^T z})^2 - \hat{\lambda} \frac{M(x_i, \hat{\zeta})}{D(x_i, \hat{\zeta})} \right], \\
 \hat{i}_{45} &= -4\hat{\lambda} \sum_{i=1}^n z_i \frac{M(x_i, \hat{\zeta}) \left[ \hat{\lambda}(x_i e^{-\hat{\beta}^T z})^2 - 1 \right]}{D(x_i, \hat{\zeta})} \left[ 1 + (\hat{\xi}-1)U(x_i, \hat{\zeta})^{-1} - \hat{\alpha}\hat{\xi}U(x_i, \hat{\zeta})^{\hat{\xi}-1} \right] \\
 &\quad - 4\hat{\alpha}\hat{\xi}\hat{\lambda}^2 \sum_{i=1}^n z_i \frac{M(x_i, \hat{\zeta})^2 U(x_i, \hat{\zeta})^{\hat{\xi}-1}}{D(x_i, \hat{\zeta})^2 U(x_i, \hat{\zeta})} [U(x_i, \hat{\zeta}) + \hat{\xi} - 1] - 4\hat{\lambda} \sum_{i=1}^n z_i (x_i e^{-\hat{\beta}^T z})^2 \\
 &\quad + 2(\hat{\xi}-1)\hat{\lambda}^2 \sum_{i=1}^n z_i \frac{M(x_i, \hat{\zeta})^2}{D(x_i, \hat{\zeta})^2 U(x_i, \hat{\zeta})^2} [1 + U(x_i, \hat{\zeta})] + 4\hat{\lambda}^2 \sum_{i=1}^n z_i \frac{M(x_i, \hat{\zeta})^2}{D(x_i, \hat{\zeta})^2}.
 \end{aligned}$$

### 2.3 Test d'ajustement pour le modèle AFT-NWR

Pour tester l'hypothèse nulle  $H_0$  selon laquelle des données proviennent d'un modèle à durée de vie accélérée dont la distribution de base est une N-Weibull-Rayleigh (AFT – NWR) :

$$H_0 : P(x_i \leq X) = F_{AFT-NWR}(x, \zeta), x > 0.$$

On propose de construire une statistique de test  $Y^2$  adapter à ce modèle. Divisons l'intervalle de temps fini  $[0, \tau]$  en  $r$  sous-intervalles  $I_j = ]a_{j-1}, a_j]$ , les fréquences théoriques  $p_j$  sont données par :

$$p_j = \int_{a_{j-1}}^{a_j} F(x) dx,$$

et les limites  $a_j$  sont calculées de sorte d'avoir des classes  $I_j$  équiprobables, ce qui nous

donne :

$$\hat{a}_j = e^{-\hat{\beta}^T z} \left[ \frac{\log \left( 1 - \exp \left\{ - \left[ -\frac{\log \left( \frac{j}{r} \right)}{\hat{\alpha}} \right]^{\frac{1}{\hat{\beta}}} \right\} \right)}{\hat{\lambda}} \right]^{\frac{1}{2}}.$$

Les valeurs des fréquences empiriques  $v_j$  obtenues en regroupant les observations dans les intervalles  $I_j$  sont données par :

$$v_j = \sum_{i=1}^n 1_{\{X_i \in I_j\}}, \quad j = \overline{1, r}.$$

## 2.4 Les éléments de la statistique du test

Afin de pouvoir calculer la statistique  $Y^2$  du test *NRR* (détaillée dans le chapitre 1, section 3.1) permettant de valider le modèle *AFT – NWR*, on a besoin de déterminer ses composantes qui sont la matrice  $\hat{G}$  et le vecteur  $\hat{L}$  du modèle étudié.

**Les éléments du vecteur  $\hat{L}$  :**

$$\begin{aligned} \frac{\partial p_j}{\partial \hat{\alpha}} &= \sum_{j=1}^r U(\hat{a}_{j-1}, \hat{\zeta})^{\hat{\xi}} F(\hat{a}_{j-1}, \hat{\zeta}) - U(\hat{a}_j, \hat{\zeta})^{\hat{\xi}} F(\hat{a}_j, \hat{\zeta}), \\ \frac{\partial p_j}{\partial \hat{\xi}} &= \hat{\alpha} \sum_{j=1}^r U(\hat{a}_{j-1}, \hat{\zeta})^{\hat{\xi}} \log [U(\hat{a}_{j-1}, \hat{\zeta})] F(\hat{a}_{j-1}, \hat{\zeta}) - \hat{\alpha} \sum_{j=1}^r U(\hat{a}_j, \hat{\zeta})^{\hat{\xi}} \log [U(\hat{a}_j, \hat{\zeta})] F(\hat{a}_j, \hat{\zeta}), \\ \frac{\partial p_j}{\partial \hat{\lambda}} &= \hat{\alpha} \hat{\xi} \sum_{j=1}^r \frac{M(\hat{a}_j, \hat{\zeta}) U(\hat{a}_j, \hat{\zeta})^{\hat{\xi}-1} F(\hat{a}_j, \hat{\zeta})}{D(\hat{a}_j, \hat{\zeta})} - \hat{\alpha} \hat{\xi} \sum_{j=1}^r \frac{M(\hat{a}_{j-1}, \hat{\zeta}) U(\hat{a}_{j-1}, \hat{\zeta})^{\hat{\xi}-1} F(\hat{a}_{j-1}, \hat{\zeta})}{D(\hat{a}_{j-1}, \hat{\zeta})}, \end{aligned}$$

$$\begin{aligned} \frac{\partial p_j}{\partial \hat{\beta}_0} &= 2\hat{\alpha} \hat{\xi} \hat{\lambda} \sum_{j=1}^r \frac{M(\hat{a}_{j-1}, \hat{\zeta}) U(\hat{a}_{j-1}, \hat{\zeta})^{\hat{\xi}-1} F(\hat{a}_{j-1}, \hat{\zeta})}{D(\hat{a}_{j-1}, \hat{\zeta})} - 2\hat{\alpha} \hat{\xi} \hat{\lambda} \sum_{j=1}^r \frac{M(\hat{a}_j, \hat{\zeta}) U(\hat{a}_j, \hat{\zeta})^{\hat{\xi}-1} F(\hat{a}_j, \hat{\zeta})}{D(\hat{a}_j, \hat{\zeta})}, \\ \frac{\partial p_j}{\partial \hat{\beta}_1} &= 2\hat{\alpha} \hat{\xi} \hat{\lambda} \sum_{j=1}^r z_i \frac{M(\hat{a}_{j-1}, \hat{\zeta}) U(\hat{a}_{j-1}, \hat{\zeta})^{\hat{\xi}-1} F(\hat{a}_{j-1}, \hat{\zeta})}{D(\hat{a}_{j-1}, \hat{\zeta})} - 2\hat{\alpha} \hat{\xi} \hat{\lambda} \sum_{j=1}^r z_i \frac{M(\hat{a}_j, \hat{\zeta}) U(\hat{a}_j, \hat{\zeta})^{\hat{\xi}-1} F(\hat{a}_j, \hat{\zeta})}{D(\hat{a}_j, \hat{\zeta})}. \end{aligned}$$

Où  $F(x, \zeta)$  est la fonction de répartition de la distribution AFT – NWR. Le vecteur  $\widehat{\mathbf{L}}_{AFT-NWR} = (\widehat{L}_1, \dots, \widehat{L}_s)^T$  est obtenu à partir de :

$$\widehat{\mathbf{L}}_{AFT-NWR}(\widehat{\zeta}) = \begin{pmatrix} \widehat{L}_1 = \sum_{j=1}^r \frac{v_j}{p_j} \frac{\partial p_j(\widehat{\zeta})}{\partial \widehat{\alpha}} \\ \widehat{L}_2 = \sum_{j=1}^r \frac{v_j}{p_j} \frac{\partial p_j(\widehat{\zeta})}{\partial \widehat{\xi}} \\ \widehat{L}_3 = \sum_{j=1}^r \frac{v_j}{p_j} \frac{\partial p_j(\widehat{\zeta})}{\partial \widehat{\lambda}} \\ \widehat{L}_4 = \sum_{j=1}^r \frac{v_j}{p_j} \frac{\partial p_j(\widehat{\zeta})}{\partial \widehat{\beta}_0} \\ \widehat{L}_5 = \sum_{j=1}^r \frac{v_j}{p_j} \frac{\partial p_j(\widehat{\zeta})}{\partial \widehat{\beta}_1} \end{pmatrix}$$

Les éléments de la matrice  $\widehat{G} = [\widehat{g}_{ll}]_{5 \times 5}$  :

$$\begin{aligned} \widehat{g}_{11} &= \widehat{i}_{11} - \sum_{j=1}^r \frac{1}{p_j} \left( \frac{\partial p_j(\widehat{\zeta})}{\partial \widehat{\alpha}} \right)^2, & \widehat{g}_{22} &= \widehat{i}_{22} - \sum_{j=1}^r \frac{1}{p_j} \left( \frac{\partial p_j(\widehat{\zeta})}{\partial \widehat{\xi}} \right)^2, \\ \widehat{g}_{33} &= \widehat{i}_{33} - \sum_{j=1}^r \frac{1}{p_j} \left( \frac{\partial p_j(\widehat{\zeta})}{\partial \widehat{\lambda}} \right)^2, & \widehat{g}_{44} &= \widehat{i}_{44} - \sum_{j=1}^r \frac{1}{p_j} \left( \frac{\partial p_j(\widehat{\zeta})}{\partial \widehat{\beta}_0} \right)^2, \\ \widehat{g}_{55} &= \widehat{i}_{55} - \sum_{j=1}^r \frac{1}{p_j} \left( \frac{\partial p_j(\widehat{\zeta})}{\partial \widehat{\beta}_1} \right)^2, & \widehat{g}_{12} &= \widehat{i}_{12} - \sum_{j=1}^r \frac{1}{p_j} \left( \frac{\partial p_j(\widehat{\zeta})}{\partial \widehat{\alpha}} \frac{\partial p_j(\widehat{\zeta})}{\partial \widehat{\xi}} \right), \\ \widehat{g}_{13} &= \widehat{i}_{13} - \sum_{j=1}^r \frac{1}{p_j} \left( \frac{\partial p_j(\widehat{\zeta})}{\partial \widehat{\alpha}} \frac{\partial p_j(\widehat{\zeta})}{\partial \widehat{\lambda}} \right), & \widehat{g}_{14} &= \widehat{i}_{14} - \sum_{j=1}^r \frac{1}{p_j} \left( \frac{\partial p_j(\widehat{\zeta})}{\partial \widehat{\alpha}} \frac{\partial p_j(\widehat{\zeta})}{\partial \widehat{\beta}_0} \right), \\ \widehat{g}_{15} &= \widehat{i}_{15} - \sum_{j=1}^r \frac{1}{p_j} \left( \frac{\partial p_j(\widehat{\zeta})}{\partial \widehat{\alpha}} \frac{\partial p_j(\widehat{\zeta})}{\partial \widehat{\beta}_1} \right), & \widehat{g}_{23} &= \widehat{i}_{23} - \sum_{j=1}^r \frac{1}{p_j} \left( \frac{\partial p_j(\widehat{\zeta})}{\partial \widehat{\xi}} \frac{\partial p_j(\widehat{\zeta})}{\partial \widehat{\lambda}} \right), \\ \widehat{g}_{24} &= \widehat{i}_{24} - \sum_{j=1}^r \frac{1}{p_j} \left( \frac{\partial p_j(\widehat{\zeta})}{\partial \widehat{\xi}} \frac{\partial p_j(\widehat{\zeta})}{\partial \widehat{\beta}_0} \right), & \widehat{g}_{34} &= \widehat{i}_{34} - \sum_{j=1}^r \frac{1}{p_j} \left( \frac{\partial p_j(\widehat{\zeta})}{\partial \widehat{\lambda}} \frac{\partial p_j(\widehat{\zeta})}{\partial \widehat{\beta}_0} \right), \\ \widehat{g}_{45} &= \widehat{i}_{45} - \sum_{j=1}^r \frac{1}{p_j} \left( \frac{\partial p_j(\widehat{\zeta})}{\partial \widehat{\beta}_0} \frac{\partial p_j(\widehat{\zeta})}{\partial \widehat{\beta}_1} \right) & j &= \overline{1, r}. \end{aligned}$$

## 2.5 Simulation numérique

### 2.5.1 Les estimateurs du maximum de vraisemblance

Étant donné que les formules analytiques des estimateurs des paramètres du modèle *AFT – NWR* ne peuvent être déterminées, nous utilisons alors les méthodes numériques pour les calculer. Pour cela, on simule 12,000 échantillons de tailles respectives  $n = 15, 50, 100, 150, 250$ , avec les valeurs initiales suivantes des paramètres ( $\alpha = 1.3, \xi = 0.7, \lambda = 2.5, \beta_0 = 1.1, \beta_1 = 3.7$ ). En utilisant l'algorithme *BBsolve* du logiciel R, on calcule les valeurs des EMV  $\hat{\alpha}, \hat{\xi}, \hat{\lambda}, \hat{\beta}_0, \hat{\beta}_1$  et leurs erreurs quadratiques moyennes (MSE), (Table 4.5)

2. CONSTRUCTION DU MODÈLE AFT-NWR

<b>n = 15</b>		<b>n = 50</b>		<b>n = 100</b>	
$\hat{\alpha}$	1.3267	$\hat{\alpha}$	1.3129	$\hat{\alpha}$	1.3077
<i>MSE</i>	$7.8940 \times 10^{-3}$	<i>MSE</i>	$6.4421 \times 10^{-4}$	<i>MSE</i>	$7.7210 \times 10^{-5}$
$\hat{\xi}$	0.7394	$\hat{\xi}$	0.7122	$\hat{\xi}$	0.7086
<i>MSE</i>	$2.2957 \times 10^{-2}$	<i>MSE</i>	$7.2998 \times 10^{-4}$	<i>MSE</i>	$1.1374 \times 10^{-4}$
$\hat{\lambda}$	2.5209	$\hat{\lambda}$	2.5126	$\hat{\lambda}$	2.5069
<i>MSE</i>	$8.9076 \times 10^{-3}$	<i>MSE</i>	$4.7196 \times 10^{-4}$	<i>MSE</i>	$1.9087 \times 10^{-4}$
$\hat{\beta}_0$	1.3316	$\hat{\beta}_0$	1.1677	$\hat{\beta}_0$	1.1249
<i>MSE</i>	$6.2974 \times 10^{-3}$	<i>MSE</i>	$8.0344 \times 10^{-4}$	<i>MSE</i>	$2.2633 \times 10^{-4}$
$\hat{\beta}_1$	3.7444	$\hat{\beta}_1$	3.7129	$\hat{\beta}_1$	3.7029
<i>MSE</i>	$5.1312 \times 10^{-3}$	<i>MSE</i>	$1.9550 \times 10^{-3}$	<i>MSE</i>	$3.8665 \times 10^{-4}$

<b>n = 150</b>		<b>n = 250</b>	
$\hat{\alpha}$	1.2996	$\hat{\alpha}$	1.2977
<i>MSE</i>	$4.5517 \times 10^{-5}$	( <i>MSE</i> )	$2.7639 \times 10^{-5}$
$\hat{\xi}$	0.6987	$\hat{\xi}$	0.6901
<i>MSE</i>	$8.2839 \times 10^{-5}$	<i>MSE</i>	$6.8397 \times 10^{-5}$
$\hat{\lambda}$	2.4936	$\hat{\lambda}$	2.4902
<i>MSE</i>	$7.8661 \times 10^{-5}$	<i>MSE</i>	$6.0148 \times 10^{-5}$
$\hat{\beta}_0$	1.1002	$\hat{\beta}_0$	1.0959
<i>MSE</i>	$1.0339 \times 10^{-4}$	<i>MSE</i>	$9.2134 \times 10^{-5}$
$\hat{\beta}_1$	3.6928	$\hat{\beta}_1$	3.6905
<i>MSE</i>	$2.5023 \times 10^{-4}$	<i>MSE</i>	$9.7026 \times 10^{-5}$

TABLE 4.5 – Les valeurs moyennes des EMV des paramètres et leurs MSE du modèle *AFT – NWR*.

Les résultats obtenus montrent que les erreurs quadratiques moyennes des estimateurs obtenus sont toutes très faibles même pour la taille des échantillons inférieure à 30. Les

résultats sont très satisfaisants. Pour conforter ces résultats, on trace dans la Figure 4.4 les valeurs absolues moyennes des paramètres estimés en fonction de la taille d'échantillon correspondante.

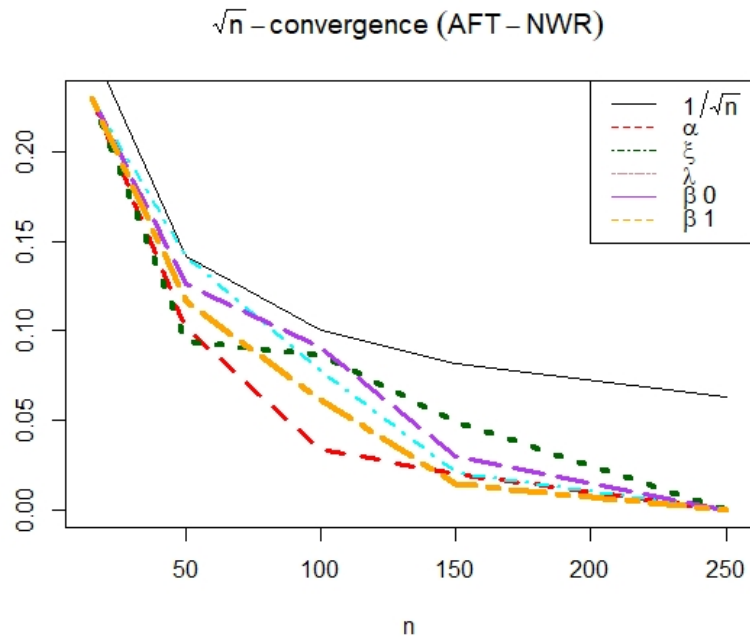


FIGURE 4.4 –  $\sqrt{n}$ -convergence des EMV de la distribution *AFT – NWR*.

On peut alors affirmer que les EMV de la distribution *AFT – NWR* obtenus sont  $\sqrt{n}$ -convergents.

### 2.5.2 La statistique du test

Pour montrer la maniabilité du test proposé et son applicabilité, on calcule dans la Table 4.6 les risques empiriques du critère de test calculé  $Y^2$  pour 12,000 échantillons de la distribution *AFT – NWR* de différentes tailles,  $n = 15$ ,  $n = 50$ ,  $n = 100$ ,  $n = 150$ ,  $n = 250$  et on les compare aux risques théoriques correspondants pour différents seuils de signification ( $\epsilon = 1\%$ ,  $5\%$ ,  $10\%$ ) :

## 2. CONSTRUCTION DU MODÈLE AFT-NWR

---

<b>n = 15</b>	<b>n = 50</b>	<b>n = 100</b>
$\epsilon = 1\%$ 0.0069	$\epsilon = 1\%$ 0.0074	$\epsilon = 1\%$ 0.0088
$\epsilon = 5\%$ 0.0340	$\epsilon = 5\%$ 0.0403	$\epsilon = 5\%$ 0.0467
$\epsilon = 10\%$ 0.0918	$\epsilon = 10\%$ 0.0933	$\epsilon = 10\%$ 0.0997

<b>n = 150</b>	<b>n = 250</b>
$\epsilon = 1\%$ 0.0112	$\epsilon = 1\%$ 0.0174
$\epsilon = 5\%$ 0.0517	$\epsilon = 5\%$ 0.0521
$\epsilon = 10\%$ 0.0955	$\epsilon = 10\%$ 0.1221

TABLE 4.6 – Comparaison des risques théoriques et des risques empiriques de la distribution *AFT – NWR* model.

À partir de la Table 4.6, nous pouvons dire que la statistique  $Y^2$  est bien adaptée pour ajuster des données au modèle *AFT – NWR*. On trace dans la figure suivante les histogrammes de la distribution des  $Y^2$  simulées par rapport à la distribution du chi-deux correspondante à  $(r - 1)$  degrés de liberté.

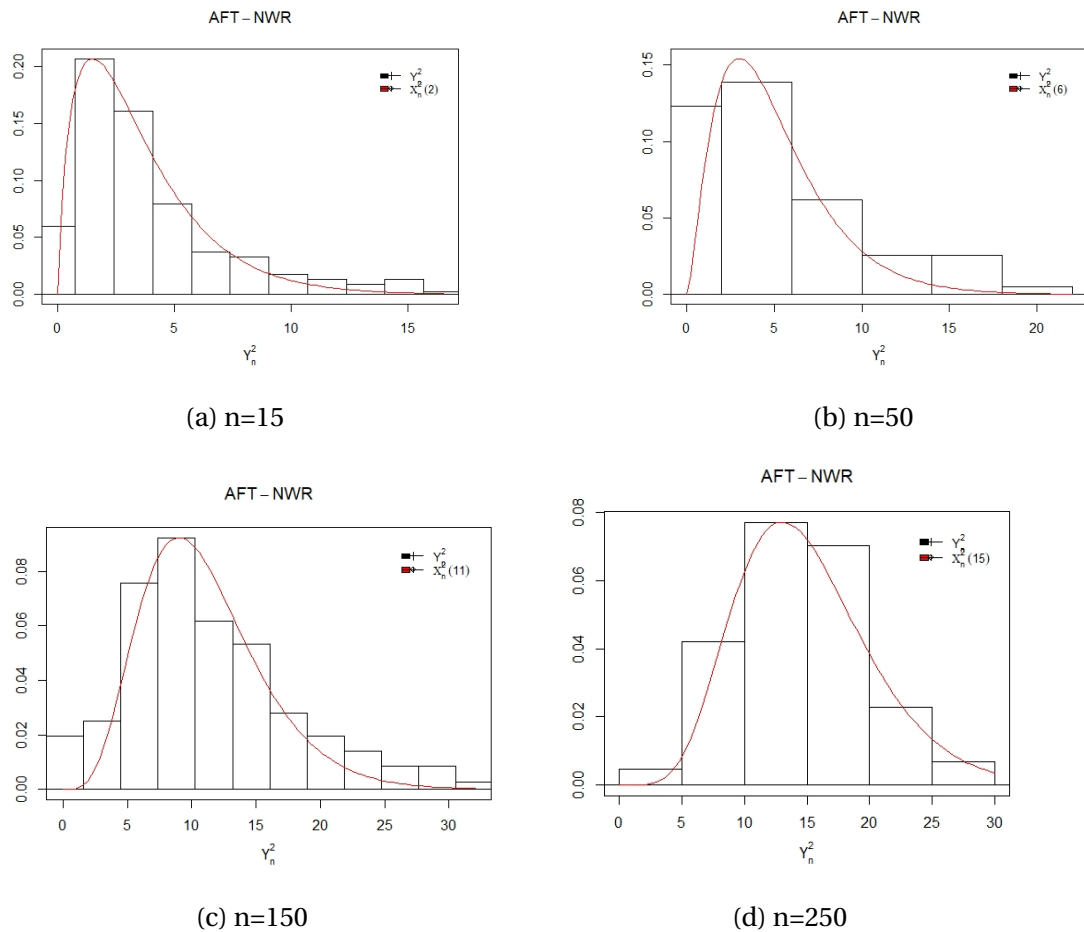


FIGURE 4.5 – Histogrammes de la distribution de  $Y^2$  et la distribution du chi-deux avec degré de liberté correspondant pour la distribution *AFT – NWR*.

## 2.6 Application aux données réelles

Une étude a été réalisée sur les données rapportées par Shin et al., (2005) sur le poids ( $x_i$ ) de 18 tumeurs exposées à différentes doses de radioactivité ( $z_i$ ) obtenues avec une technique médicale particulière (images scintigraphiques), (Table 4.7).

$x_i$	0.04, 0.06, 0.07, 0.09, 0.11, 0.11, 0.21, 0.23, 0.23, 0.27, 0.27, 0.33, 0.36, 0.41, 0.49, 0.51, 0.60, 0.70
$z_i$	0.779, 2.078, 1.299, 1.818, 3.636, 1.948, 7.013, 8.052, 6.494, 10, 4.675, 5.974, 8.312, 7.273, 6.494, 11.169, 10.390, 9.351

TABLE 4.7 – Le poids de 18 tumeurs exposés à différentes doses de radioactivité.

Nous supposons l'hypothèse nulle  $H_0$  que nous pouvons ajuster ces observations par la distribution  $AFT - NWR$  et nous traçons dans la Figure 4.6 les PP et QQ plot de ces données par rapport à ceux de la distribution  $AFT - NWR$

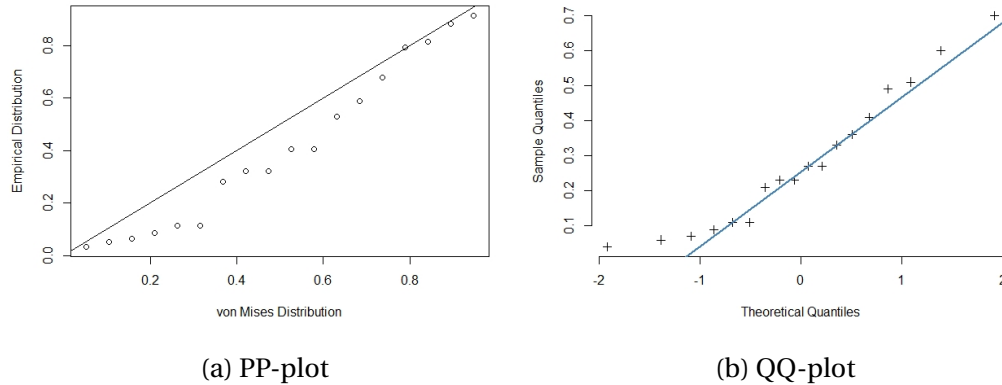


FIGURE 4.6 – PP et QQ plot du poids de 18 tumeurs exposés à différentes doses de radioactivité.

On peut observer que la distribution du poids des tumeurs a presque la même forme que les quantiles et les probabilités théoriques du modèle  $AFT - NWR$ .

Pour confirmer l'hypothèse nulle  $H_0$ , on calcule le vecteur des EMV du modèle  $AFT - NWR$  :

$$\hat{\zeta} = (\hat{\alpha} = 3.5474, \hat{\xi} = 0.5292, \hat{\lambda} = 0.1217, \hat{\beta}_0 = -3.0324, \hat{\beta}_1 = 0.18873)^T.$$

En regroupant les données observées en  $r = 4$  intervalles ( $I_j$ ), on calcule les limites  $\hat{a}_j$  de ces intervalles, les fréquences observées  $\hat{v}_j$  et les probabilités attendues  $p_j$  correspondantes :

$\hat{a}_j$	0.1626	0.2102	0.2599	100.70
$\hat{v}_j$	6	1	2	9
$p_j$	0.25	0.25	0.25	0.25

TABLE 4.8 – Valeurs de  $\hat{a}_j$ ,  $\hat{v}_j$  et  $p_j$ .

La matrice d'information de *Fisher* estimée  $\hat{I}(\hat{\zeta})$  est alors égale à :

$$\hat{I}(\hat{\zeta}) = \begin{pmatrix} 6.25 & 10.0875 & -11.2254 & 11.2254 & 51.8609 \\ 10.0875 & 4.2617 & -6.0116 & 6.0116 & 21.2706 \\ -11.2254 & -6.0116 & 9.8023 & -5.6423 & -6.3847 \\ 11.2254 & 6.0116 & -5.6423 & 3.0340 & -870.4102 \\ 51.8609 & 21.2706 & -6.3847 & -870.4102 & 15.8255 \end{pmatrix}.$$

Par conséquent, nous obtenons la valeur du critère du test statistique,  $Y^2 = 4.7651$ . Nous la comparons à la distribution de chi-deux  $\chi^2_\epsilon$  à  $(4 - 1) = 3$  degrés de liberté pour différents niveaux de signification  $\epsilon = 1\%$ ,  $\epsilon = 5\%$  et  $\epsilon = 10\%$  :

$$Y^2 < \chi^2_{1\%}(3) = 11.3448,$$

$$Y^2 < \chi^2_{5\%}(3) = 7.8147,$$

$$Y^2 < \chi^2_{10\%}(3) = 6.2513.$$

Les résultats obtenus affirment que la distribution de ces observations peut être modélisée par la distribution *AFT – NWR*. Par conséquent, nous pouvons conclure que la distribution *AFT – NWR* modélise efficacement les données médicales.

# Conclusion

Après avoir présenté les différentes méthodes de généralisation de la distribution Weibull et les approches existantes, on s'est intéressé particulièrement à la Nouvelle Famille Weibull-G. Cette étude nous a permis de mieux comprendre les propriétés et les applications de ces distributions dans divers domaines tels que la fiabilité, l'analyse de survie, les études économiques et autres. On s'est proposé alors de construire de nouveaux modèles, le modèle N-Weibull-Weibull (NWW) et le modèle N-Weibull-Rayleigh (NWR), et leurs modèles à temps de vie accéléré correspondants. Ces distributions sont capables de mieux décrire une plus large variété de données. L'analyse statistique complète réalisée pour chaque modèle a permis d'évaluer leur adéquation aux données réelles. Des tests d'ajustement spécifiques ont été développés et les estimations des paramètres inconnus ont été effectuées en utilisant la méthode du maximum de vraisemblance. Des simulations numériques approfondies ont été menées pour valider la performance des estimateurs et des critères de test proposés.

En conclusion, on souhaite que ce modeste travail de recherche a contribué un peu à la théorie et à la pratique de la statistique. Les méthodes d'estimation et de tests développées ont démontré leur efficacité, offrant ainsi des outils statistiques avancés pour les chercheurs et les différents utilisateurs.

## Conclusion

---

Ces travaux ouvrent également des perspectives de recherche prometteuses pour l'extension de ces approches à d'autres distributions, pour l'exploration de nouvelles applications dans des domaines variés et offre des opportunités pour de futures investigations et développements dans ce domaine.

# Bibliographie

- [1] Alexander, C., Cordeiro, G.M., Ortega, E.M.M. and Sarabia, J.M (2012). *Generalized beta generated distributions*. Computational Statistics and Data Analysis 56, 1880-1897.
- [2] Al-Noor, N.H., Assi, N.K. (2020). *Rayleigh-Rayleigh Distribution : Properties and Applications*. Journal of Physics : Conference Series. 1591(1).
- [3] Alzaatreh, A., Lee, C., Famoye, F. (2013). *A new method for generating families of continuous distributions*. Metron. 71(1), 63-79.
- [4] Alzaatreh, A., Famoye, F., Lee, C. (2014). *The gamma-normal distribution. Properties and applications*. Comput. Stat. Data Anal. 69(1), 67-80.
- [5] Amini, M., MirMostafaei, S.M.T.K. and Ahmadi J. (2014). *Log-gamma-generated families of distributions*. Statistics 48, 913-932.
- [6] Azzalini, A. (1985). *A class of distributions which includes the normal ones*. Scand. J. Stat. 12, 171-178.
- [7] Bagdonavičius, V. (1978). *Testing the hypothesis of the additive accumulation of damages*. Probability Theory and its Application, 23(2), 403–408.

- [8] Bagdonavičius, V., Levuliene, R., and Nikulin, M. (2012). Chi-squared goodness-of-fit tests for parametric accelerated failure time models. *Communications in Statistics - Theory and Methods*.
- [9] Bagdonavicius, V. and Nikulin, M. (2002). *Accelerated Life Models*. Chapman and Hall-CRC, Boca Raton.
- [10] Bagdonavicius, V. and Nikulin, M. (2011). *Chi-square goodness-of-fit test for right censored data*. *Int.J.Appl.Math.Stat.* 1(11), 30-50.
- [11] Bagdonavicius, V., Nikulin, M., and Tahir, R. (2010c). Chi-squared goodness-of-fit test for generalized weibull distribution when data are right censored. In 3rd International conference on accelerated life testing, reliability-based analysis and design. 223–230, ClermontFerrand, France.
- [12] Balakrishnan, N., Kateri, M. (2008). *On the maximum likelihood estimation of parameters of Weibull distribution based on complete and censored data*. *Statistics and Probability Letters*. 78(17), 2971-2975.
- [13] Bourguignon, M., Silva, R.B. and Cordeiro, G.M. (2014). *The Weibull–G family of probability distributions*. *Journal of Data Science* 12, 53-68.
- [14] Burr, I.W. (1942). *Cumulative frequency functions*. *Ann. Math. Stat.* 13, 215-232.
- [15] Cordeiro, G. M. and de Castro, M. (2011). *A new family of generalized distributions*. *Journal of Statistical Computation and Simulation*, 81, 883-898.
- [16] Cordeiro, G.M., Silva, G.O. and Ortega, E.M.M (2012). *The beta extended Weibull distribution*. *Journal of Probability and Statistical Science (Taiwan)* 10, 15-40.
- [17] David, G., Kleinbaum (1995). *Survival analysis-Statistics in the health sciences*. 3e édition Springer, 3-14.
- [18] Dodge, Y., Melfi, G. (2008). *Premiers pas en simulation*. Springer-Verlag France, Springer Paris Berlin Heidelberg New York. 978-2-287-79493-3.

- 
- [19] Drost, F. (1988). Asymptotics for generalized chi-squared goodness-of-fit tests. CWI Tracts, Amsterdam, 48. Centre for Mathematics and Computer Sciences. 3, 23, 29.
- [20] Eugene, N., Lee, C., Famoye, F. (2002). *The beta-normal distribution and its applications*. Commun. Stat. Theo. Meth.31(4), 497-512.
- [21] Ferreira, J.T.A.S., Steel, M.F.J. (2006). *A constructive representation of univariate skewed distributions*. J.Am. Stat. Assoc. 101(474), 823-829.
- [22] Freimer, M., Kollia, G., Mudholkar, G.S., Lin, C.T. (1988). *A study of the generalized Tukey lambda family*. Commun. Stat. Theo. Meth. 17, 3547-3567.
- [23] Fry, T.R.L. (1993). *Univariate and multivariate Burr distributions : a survey*. Pak. J. Stat. Ser. A 9, 1-24.
- [24] Gauss, C.F. (1809). *Theoria Motus Corporum Coelestium*. Perthes u. Besser, Hamburg. Lib. 2, Sec. III. 205-224.
- [25] Gauss, C.F. (1816). *Bestimmung der genauigkeit der beobachtungen*. Zeitschrift Astronomi. 1, 185-197.
- [26] Goldfarb Bernard Catherine Pardoux (2011). *Introduction à la méthode statistique*. 6e édition Dunod, Paris.
- [27] Goual H., Seddik Ameer (2016). *A modified chi-squared goodness of fit test for the Kumaraswamy generalized inverse Weibull distribution and its applications*. Journal of Statistics : Advances in Theory and Applications, 16(2), 275-305.
- [28] Gross A.J., Clark V.A. (1975). *Survival Distributions : Reliability Applications in the Biometrical Sciences*, John Wiley, New York, USA.
- [29] Gupta, R. C., Gupta, P. I. and Gupta, R. D. (1998). *Modeling failure time data by Lehmann alternatives*. Communications in statistics-Theory and Methods 27, 887-904.

- [30] He, W., Xiong, J. and Yi, G.Y., (2012). *SIMEX R Package for Accelerated Failure Time Models with Covariate Measurement Error*. Journal of Statistical Software. 46(1), 1-14. <http://doi.org/10.18637/jss.v046.c01>.
- [31] <https://data.world/datasets/data>.
- [32] <https://www.donneesmondiales.com>.
- [33] Johnson, N.L. (1949). *Systems of frequency curves generated by methods of translation*. Biometrika 36, 149-176.
- [34] Johnson, N.L., Kotz, S., Balakrishnan, N. (1994). *Continuous Univariate Distributions*, vol. 1, 2nd edn. Wiley, New York.
- [35] Jones, M.C (2004). *Families of distributions arising from distributions of order statistics*. Test 13(1), 1-43.
- [36] Karian, Z.A., Dudewicz, E. (2000). *Fitting Statistical Distributions The Generalized Lambda Distribution and Generalized Bootstrap Methods*. Chapman Hall/CRC Press, Boca Raton.
- [37] Kotz S., Nadarajah S. (2000). *Extreme Value Distributions. Theory and Applications*. Imper. Coll. Press. London.
- [38] Kumaraswamy, P. (1980). *Generalized probability density-function for double-bounded random-processes*. Journal of Hydrology, 46, 79-88.
- [39] Lafaye, P. de Micheaux, Drouilhet, R., Liquet, P. (2011). *Le logiciel R, Maitriser le langage, Effectuer des analyses statistiques*. Springer-Verlag France. Springer Paris Berlin Heidelberg New York. 978-2-8178-0114-8.
- [40] Mathers, B.M., Degenhardt, L., Bucello, C., Lemon, J., Wiessing, L., Hickman, M. (2013). *Mortality among people who inject drugs : a systematic review and meta-analysis*. Bull World Health Organ. 91(2), 102-123.
- [41] Meeker, W. Q., Escobar, L.A., Pascual, F.G. (1998). *Statistical Methods for Reliability Data, 2nd Edition*, John Wiley and Sons Inc.

- 
- [42] Mudholkar, G.S., Srivastava, D.K. and Freimer, M. (1995). *The exponentiated Weibull family : A reanalysis of the bus-motor failure data*. Technometrics 37, 436-445.
- [43] Nadarajah, (2006). *The exponentiated Gumbel distribution with climate application*, Environmetrics. 17, 13-23.
- [44] Nadarajah, S., Cordeiro, G.M., Ortega, E.M.M. (2015). *The Zografos-Balakrishnan-G family of distributions : Mathematical properties and applications*. comm. in stat. 44(1), 186-215.
- [45] Nadarajah, S., Kotz, S. (2004). *The beta Gumbel distribution*. Math. Probl. Eng. 4, 323-332.
- [46] Nadarajah, S., Kotz, S. (2005). *The beta exponential distribution*. Reliab. Eng. Syst. Saf. 91(6), 689-697.
- [47] Nadarajah, S., Kotz, S. (2006). *The exponentiated type distributions*. Acta Applicandae Mathematica 92, 97-111.
- [48] Nikulin, M.S. (1973a). *Chi-square Test For Continuous distributions with Shift and Scale Parameters*. Theory of Probability and Its Applications. 18(3), 559-568.
- [49] Nikulin, M.S. (1973b). *On a Chi-square test for continuous distributions*. Theory of Probability and its Applications. 18(3), 638 – 639.
- [50] Pearson, K. (1895). *Contributions to the mathematical theory of evolution. II. Skew variation in homogeneous material*. Philos. Trans. Royal Soc. Lond. A 186, 343-414.
- [51] Proschan, F (1963). *Theoretical explanation of observed decreasing failure rate*. Technometrics 5(3) : 375-383.
- [52] Ramberg, J.S., Schmeiser, B.W. (1972). *An approximate method for generating symmetric random variables*. Commun. Assoc. Comput. Mach. 15, 987-990.
- [53] Ramberg, J.S., Schmeiser, B.W. (1974). *An approximate method for generating asymmetric random variables*. Commun. Assoc. Comput. Mach. 17, 78 – 82.

- [54] Ramberg, J.S., Tadikamalla, P.R., Dudewicz, E.J., Mykytka, E.F. (1979). *A probability distribution and its uses in fitting data*. Technometrics 21, 201-214.
- [55] Rao, K.C., Robson, D.S. (1974). *A chi-square statistic for goodness-of-fit tests within the exponential family*. Communications in Statistics. 3(12), 1139-1153.
- [56] Ravi, V., Gilbert, P.D. (2009). *BB : An R Package for Solving a Large System of Nonlinear Equations and for Optimizing a High-Dimensional Nonlinear Objective Function*. Journal of Statistical Software.32(4) , 1–26.
- [57] Ristic, M.M., Balakrishnan, N. (2011). *The gamma exponentiated exponential distribution*. Journal of Statistical Computation and Simulation. 82(8), 1191-1206.
- [58] Ristic, M. M. and Balakrishnan, N. (2012). *The gamma-exponentiated exponential distribution*. Journal of Statistical Computation and Simulation 82, 1191-1206.
- [59] Shin, JH., Chung, JK., Kang, JH., Lee, YJ., Kim, KI., Jeong, JM., Lee, DS., (2005). *Noninvasive imaging for monitoring of viable cancer cells using a dual-imaging reporter gene*. The Journal of Nuclear Medicine, 45(12), 2109-2115.
- [60] Tahir, M.H., M. Zubair, M. Mansoor, Gauss M. Cordeiro, Morad Alizadeh and G. G. Hamedani (2016). *A new Weibull-G family of distributions*. Hacettepe Journal of Mathematics and Statistics. 45(2), 629 – 647.
- [61] Tilbi, D., Seddik-Ameur, N. (2017). *Chi-squared goodness-of-fit tests for the generalized Rayleigh distribution*. Journal of Statistical Theory and Practice. 11(4), 594-603.
- [62] Treidi, W., Seddik-Ameur, N. (2016). *NRR statistic for the extension Weibull distribution*. Global Journal of Pure and Applied Mathematics. 12(4), 2809-2818.
- [63] Tukey, J.W. (1960). *The practical relationship between the common transformations of percentages of counts and amounts*. Technical Report 36, Statistical Techniques Research Group, Princeton University, Princeton, NJ.
- [64] Van der Vaart, A. (1998). Asymptotic Statistics. Cambridge, UK.

- [65] Voda, V.G. (1972). On the Inverse Rayleigh Distribution Random Variables. Report in Statistical Applied Research JUSE. 19, 13-21.
- [66] Yousof, M.H., Masoom, M.A., Goual, H., Ibrahim, M. (2021). *A new reciprocal Rayleigh extension : properties, copulas, different methods of estimation and a modified right-censored test for validation*. Statistics in Transition New Series. 22(3), 99-121.
- [67] Zografos, K., Balakrishnan, N. (2009). *On families of beta- and generalized gamma generated distributions and associated inference*. Stat. Meth. 6, 344-362.