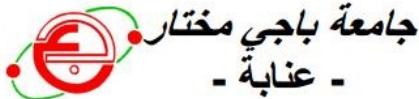


وزارة التعليم العالي والبحث العلمي

BADJI MOKHTAR -ANNABA
UNIVERSITY
UNIVERSITE BADJI MOKHTAR
ANNABA



Faculté des Sciences
Département de Mathématiques

THESE

Présentée en vue de l'obtention du diplôme de
DOCTORAT 3^{ème} CYCLE EN MATHEMATIQUES
Option : Statistique

Etude d'un modèle Inverse Weibull généralisé à temps de vie accéléré

Par
Mlle GOUAL HAFIDA

Sous la direction de

Pr SEDDIK AMEUR NACIRA

Devant le jury

PRESIDENT	REMITA Med Riad	M.C.A	U.B.M. ANNABA
RAPPORTEUR	SEDDIK AMEUR Nacira	Pr	U.B.M. ANNABA
EXAMINATEUR	CHADLI Assia	M.C.A	U.B.M. ANNABA
EXAMINATEUR	NIKULIN Mikhail	Pr Emérite	U.V.S. Bordeaux 2
EXAMINATEUR	MESSACI Fatiha	Pr	U.M. Constantine
EXAMINATEUR	NEMOUCHI Nahima	Pr	U.M. Constantine

Année Universitaire 2013 / 2014

REMERCIEMENTS

A l'issue de la rédaction de cette recherche, je suis convaincue que la thèse est loin d'être un travail individuel. En effet, je n'aurai jamais pu réaliser ce travail doctoral sans le soutien d'un grand nombre de personnes dont la générosité, la bonne humeur et l'intérêt manifestés à l'égard de ma recherche m'ont permis de progresser dans cette phase délicate de l'apprenti-chercheur.

Je voudrais remercier chaleureusement et en premier lieu **le professeur Seddik ameur Nacira**, ma directrice de thèse qui n'a pas simplement accepté de diriger ma thèse mais, également m'a transmis la passion de la recherche mathématique et qui n'a pas cessé de m'encourager et de me soutenir durant ces années. J'ai pu apprécier non seulement sa dimension mathématique, mais aussi son non moins importante dimension humaine. Pour sa façon de faire, de laisser faire et parfois de ne pas faire, qui m'a surprise si souvent. J'en profite pour vous exprimer ici ma plus profonde gratitude. Celui de ses concepts qui restera mon favori : "faire au moins une chose par jour" : conseil salutaire s'il en est un.

Un remerciement tout particulier au **professeur Nikulin Mikhail**, qui m'a accueillie au Laboratoire de l'université de Bordeaux 2 et pour avoir accepté de faire partie de mon jury.

Je remercie les rapporteurs, **les professeurs Messaci Fatiha et Nemouchi Nahima**, dont la gentillesse n'a d'égale que leurs immenses compétences mathématiques, d'avoir accepté la lourde tâche de lire cette thèse jusque dans ses détails, et d'en tirer un compte-rendu.

Je remercie très sincèrement **Dr Chadli Assia**, une mathématicienne dont la carrière est impressionnante, d'avoir accepté d'examiner cette thèse

avec attention et de faire partie de mon jury.

Je souhaite exprimer ma reconnaissance au doyen de la faculté des sciences de l'université Badji Mokhtar annaba, **Dr. Remita Med Riad**, pour son enthousiasme et ses encouragements continus à tous les mathématiciens de notre université, de m'avoir fait l'honneur de présider le jury.

Je remercie également tous les membres du département de mathématiques de l'université Badji Mokhtar annaba, en particulier ceux du laboratoire de probabilités et statistiques **LaPS**, parce que c'est aussi grâce à eux que ce laboratoire est si agréable et si stimulant.

Je profite aussi de cette occasion pour remercier tous les professeurs que j'ai eu dans ma scolarité et qui m'ont donné l'envie d'apprendre et d'apprendre toujours plus.

A vous qui ne lirez de cette thèse que ces remerciements et quelques lignes de l'introduction, les mots les plus simples étant les plus forts, j'adresse toute mon affection à ma famille, et en particulier à ma **mère** qui m'a fait comprendre que la vie n'est pas faite que de problèmes, des problèmes qu'on pourrait résoudre grâce à des formules mathématiques, à mon **père** qui m'a toujours soutenue et poussée en avant, faisant fin de mes doutes et mes objections.

Je remercie bien sur mes soeurs **Naima** et **Saiko**, mes frères **Kamel** et **Midou**, de m'avoir supportée toutes ces années. Leur confiance, leur tendresse, leur amour me portent et me guident tous les jours. Merci pour avoir fait de moi ce que je suis aujourd'hui. Est-ce un bon endroit pour dire ce genre de choses ? Je n'en connais en tous cas pas de mauvais. Je vous aime !

A mes amies **Nafissa** et **Manel** qui, avec cette question récurrente, "quand est-ce que tu la soutiens cette thèse ?", bien qu'angoissante en période fréquente de doutes, m'ont permis de ne jamais dévier de mon objectif final.

Je finis ces remerciements par la personne qui subit de plein fouet mon stress naturel et les fluctuations d'humeur inhérentes à une thèse, **Lamia** : merci de me remettre sur le droit chemin quand il le faut. Merci pour ton soutien tout au long de ma thèse, surtout dans la dernière ligne droite où tu as su me rassurer et me redonner la confiance qu'il me fallait !.

Abstract

Accelerated failure time models deals with random variables which distribution survival time depends on explanatory variables that can represent, for example medication doses, temperature, humidity, or stress. The well known models are parametric AFT models (accelerated failure time), developed by Bagdonavičius and Nikulin (2002). These AFT models stand out models called proportional hazards (proportional hazard or PH) which are characteristic of semi-parametric modellings, in particular the most used of them, the Cox models.

In this work, we construct an AFT-GIW model, for which the baseline distribution is a new three-parameters distribution, proposed by Gusmao et al. (2009) and called generalized inverse Weibull. After studying the characteristics of the model, we construct a modified Chi-squared goodness-of-fit test based on the NRR (Nikulin-Rao-Robson) statistic in both of censored and complete data. We use the maximum likelihood estimators based on initial non grouped data. Then, we conduct numerical simulations to corroborate the results. The study of different samples from real data from reliability and survival analysis, showed the applicability of this model in various fields.

Keywords : Accelerated failure time model, censored data, Chi-squared type goodness-of-fit test, generalized inverse Weibull distribution, NRR statistic, survival analysis.

Résumé

Les modèles à durée de vie accélérée traite des variables aléatoires dont la distribution des temps de survie dépend de variables explicatives qui peuvent représenter par exemple les doses de médicaments, la température, l'humidité, ou le stress. Les modèles les plus connus sont les modèles paramétriques AFT(accelerated failure time), développés par Bagdonavicius et Nikulin (2002). Ces modèles AFT se distinguent des modèles dits à risques proportionnels (proportional hazard ou PH) qui sont caractéristiques des modélisations semi-paramétriques et notamment de la plus utilisée d'entre elles, le modèle de Cox.

Dans ce travail, nous construisons un modèle AFT-IWG, dont la distribution de base est une distribution récente à trois paramètres, proposée par Gusmao et al. (2009), appelée inverse Weibull généralisée. Après l'étude des caractéristiques du modèle, nous construisons pour celui-ci, un test d'ajustement du type du Chi-deux modifié, basé sur la NRR (Nikulin-Rao-Robson) statistique dans le cas de données censurées et de données complètes. Nous utilisons les estimateurs du maximum de vraisemblance basés sur les données initiales non regroupées. Ensuite, nous menons des simulations numériques pour corroborer les résultats obtenus. L'étude de différents échantillons issus de données réelles de fiabilité et d'analyse de survie montrent l'applicabilité de ce modèle dans différents domaines.

Mots-clés : Analyse de survie, distribution inverse Weibull généralisée, données censurées, modèle de temps de défaillance accéléré, statistique NRR, tests d'ajustement du type du Chi-deux.

ملخص

نموذج وقت الإخفاق المتسارع، يتعامل مع المتغيرات العشوائية التي توزع الوقت لديها يعتمد على المتغيرات التفسيرية التي يمكن أن تمثل، على سبيل المثال، جرارات الدواء، درجة الحرارة، الرطوبة، أو الإجهاد. النماذج المعروفة هي نماذج بارامترية AFT (المتسارع الوقت فشل)، التي وضعها Bagdonavicius and Nikulin (2002)، هذه النماذج AFT تبرز نماذج المخاطر النسبية تسمى(الخطر النسبي او PH). التي هي سمة من نمذجة شبه بارامترية ولا سيما الأكثر استخداماً منهم، نماذج Cox.

في هذا العمل، نقوم بتصميم نموذج AFT-GIW، الذي توزيعه القاعدي هو توزيع جديد بثلاثة بارامترات ، التي اقترحها Gusmao et al (2009)، تسمى generalized inverse Weibull . بعد دراسة خصائص النموذج، نقوم بإنشاء، اختبار الظبط المعدل كا يستند على إحصائية- NRR (Nikulin-Rao- Robson) في حالة البيانات المراقبة والبيانات الكاملة. نستخدم مقدرات احتمال الحد الأقصى بناء على بيانات أولية غير مجمعة. ثم، نعمل على إجراء المحاكاة العددية لتأكيد صحة النتائج. أظهرت دراسة عينات فعلية مختلفة من بيانات فعلية وبيانات تحليل البقاء على قيد الحياة إمكانية تطبيق هذا النموذج في مختلف المجالات.

الكلمات الرئيسية:

نموذج وقت الإخفاق المتسارع AFT ، generalized inverse Weibull ، اختبار الظبط المعدل كا ، NRR الاختبار ، تحليل البقاء على قيد الحياة ، البيانات المراقبة.

Table des matières

Introduction	1
1 Présentation du modèle inverse Weibull généralisé (IWG)	5
1.1 Quelques définitions	5
1.1.1 Censure	5
1.1.2 Fonction de vraisemblance	7
1.1.3 Estimateur de Kaplan-Meier	7
1.2 La distribution inverse Weibull (IW)	9
1.3 La distribution inverse Weibull généralisée (IWG)	13
1.4 Fonction génératrice des moments	18
1.5 Estimation du maximum de vraisemblance en cas de données complètes	19
1.5.1 Estimation du maximum de vraisemblance	19
1.5.2 Matrice d'information de Fisher	21
1.6 Estimation du maximum de vraisemblance en cas de données censurées	22
2 Construction du modèle AFT-IWG	25
2.1 Introduction des modèles AFT	26
2.1.1 Définition du modèle AFT	28
2.1.2 Plans d'expérience	30
2.2 Construction du modèle AFT-IWG	32
2.3 Estimation du maximum de vraisemblance en cas de données censurées	34
2.4 Estimation et intervalle de confiance de la fonction de survie . .	37
2.4.1 Estimation de la fonction de survie	37
2.4.2 Intervalle de confiance de la fonction de survie	37

3 Tests d'ajustement pour données censurées	41
3.1 Tests d'ajustement pour des données censurées à droite	42
3.1.1 Test de Habib et Thomas	42
Test de Habib et Thomas pour la distribution IWG	45
3.1.2 Test de Bagdonavičius et Nikulin	47
3.2 Application du test NRR pour la distribution IWG	56
3.2.1 Calcul de la matrice \hat{W}	58
3.2.2 Calcul de la matrice de Fisher estimée \hat{I}	58
3.3 Test d'ajustement pour le modèle AFT inverse Weibull généralisé en cas de données censurées	61
3.3.1 Choix des intervalles	62
3.3.2 Calcul de la matrice \hat{W}	62
3.3.3 Matrice d'information de Fisher estimée \hat{I}	64
4 Simulations et applications numériques	73
4.1 Simulation de la distribution IWG	73
4.1.1 Estimateurs du maximum de vraisemblance des paramètres de la distribution IWG	73
4.1.2 Distribution de la statistique Y_n^2 pour la distribution IWG	75
4.1.3 Test Y_n^2	78
4.2 Simulation du modèle AFT-IWG	85
4.2.1 Estimateurs du maximum de vraisemblance des paramètres du modèle AFT-IWG	85
4.2.2 Distribution de la statistique Y_n^2	87
4.2.3 La statistique du test Y_n^2 pour le modèle AFT-IWG	88
4.3 Applications à la distribution IWG	89
4.3.1 Exemple de données de survie	89
4.3.2 Exemple de fiabilité	90
4.4 Applications au modèle AFT-IWG	92
4.4.1 Exemples de données de survie	92
4.4.2 Exemple de données de fiabilité	97
Conclusion et perspectives	100
Bibliographie	103

Introduction

L'analyse des données de survie est l'étude du délai de la survenue d'un événement précis comme par exemple l'apparition d'une maladie en analyse de survie, ou bien la panne d'une machine (durée de fonctionnement d'une machine, en fiabilité), ou encore la survenue d'un sinistre en actuariat. Les statistiques développent des techniques et des méthodes qui permettent d'analyser les données issues de l'observation, afin de cerner les caractéristiques de la population concernée et d'identifier un modèle capable d'engendrer ces données. Les modèles utilisés peuvent être paramétriques, non paramétriques ou semi-paramétriques.

Il arrive souvent que les fonctions d'intérêt peuvent dépendre de covariables comme par exemple les doses de médicaments, la température, l'humidité, ou le stress. Celles-ci sont susceptibles d'influencer la survie et donc de modifier les fonctions de risque. En analyse de survie, les modèles largement répandus pour analyser ce type de données, sont les modèles à hasards proportionnels (PH proportional hazards) ou les modèles de Cox. Tandis que dans l'industrie, ce sont les modèles de taux de pannes accéléré AFT (accelerated failure time) qui sont les plus utilisés. Ces modèles peuvent être aussi bien paramétriques, semi paramétriques ou non paramétriques (voir Bagdonavičius (1978), et Bagdonavičius et Nikulin (2002)).

Les modèles AFT ou modèles à temps de vie accéléré ou tests de vie accélérés (accelerated life testing) est une des approches les plus communes pour obtenir la loi de fiabilité ou autre caractéristiques comportementales (taux de défaillances, temps de survie, etc...) des produits (systèmes ou composants) dans des délais courts. Ils consistent à réduire les durées de vie des produits par l'accélération des dégradations provoquant les défaillances. Pour cela les niveaux de stress subis par le produit sont augmentés afin d'ob-

tenir des données de vie plus rapidement, qui seront utilisées pour estimer la fiabilité dans les conditions normales de fonctionnement. Ils permettent d'améliorer les performances des systèmes et d'optimiser les stratégies de maintenance.

Quelques modèles AFT classiques ont été traités dans la littérature statistique tels que le modèle AFT- exponentiel dont la distribution de base est une loi exponentielle, AFT- Weibull, AFT- lognormal et AFT- loglogistic (voir Bagdonavičius et Nikulin (2002), Lawless (2003), Meeker et Escobar (1998), Nelson (1990)). Plus tard, Bagdonavičius et al. (2004) se sont intéressés à l'estimation dans le cas du modèle AFT-Weibull où la distribution de base est une distribution Weibull à deux paramètres. En 2011, Bagdonavičius et Nikulin (2011, 2011a), et Bagdonavičius et al. ont proposé des tests d'ajustement du type du Chi-deux pour les modèles sus-cités.

Notre étude consiste, d'abord, à construire un modèle AFT-IWG, dont la fonction de base est une distribution inverse Weibull généralisée à trois paramètres, proposée récemment par Gusmao et al. (2009). C'est une distribution dont la forme du risque instantané (taux de hasard) peut être constante, monotone croissante, monotone décroissante et en forme \cup ou en forme de \cap . En plus de sa flexibilité et de ses interprétations physiques, elle peut représenter n'importe quelle situation dans une analyse de durées de vie, que ce soit en fiabilité ou en analyse de survie.

Après l'étude des caractéristiques du modèle, nous construisons un test d'ajustement du type du Chi-deux modifié basé sur la statistique NRR (Nikulin-Rao-Robson) proposée par Bagdonavičius et Nikulin (2011), Bagdonavičius et al. (2011) que nous avons adaptée pour notre modèle dans le cas de données censurées. La distribution de la statistique de ce test, basée sur les estimateurs du maximum de vraisemblance sur les données initiales non groupées, tend vers une distribution du Chi-deux à k degrés de liberté, où k représente le nombre de classes choisi. Nous donnons la forme explicite de tous les éléments de la statistique de test. Ensuite, nous menons des simulations numériques pour corroborer les résultats obtenus. L'étude de différents échantillons issus de données réelles de fiabilité et d'analyse de survie montrent l'applicabilité de ce modèle dans différents domaines.

D'autre part, étant donné que la distribution inverse Weibull généralisée

est récente, les propriétés mathématiques, les moments et la mixture de deux distributions IWG ont été traités (Gusmao (2009)), mais la validation du modèle n'a pas encore été investie. Nous avons alors, développé un test d'ajustement du Chi-deux pour ce modèle IWG dans le cas de données censurées permettant ainsi de vérifier si une série d'observation peut être modéliser par cette distribution très flexible.

Le manuscrit s'articule autour de quatre chapitres.

Dans le premier chapitre, nous étudions les propriétés statistiques et mathématiques de la distribution inverse Weibull généralisée, à trois paramètres, proposée par Gusmao et al. (2009). Nous y exposons la fonction de fiabilité, la fonction risque, les estimateurs du maximum de vraisemblance des paramètres inconnus sur les données non regroupées et le calcul de la matrice d'information de Fisher ; et ceci en présence de censure aléatoire droite.

Dans le deuxième chapitre, nous proposons la construction d'un modèle de durée de vie, à temps de vie accéléré, c'est le modèle AFT-IWG dont la distribution de base est celle de l'inverse Weibull généralisée. Nous présentons une étude exhaustive du modèle, fonction de fiabilité, fonction risque selon toutes ses formes décroissante, constante et croissante. Nous utilisons la méthode du maximum de vraisemblance sur les données initiales pour estimer les paramètres inconnus du modèle, en tenant compte de la censure.

Quand au troisième chapitre, il est consacré à la construction de tests d'ajustement du type du Chi-deux modifié pour le modèle proposé et pour le modèle inverse Weibull généralisé. Nous utilisons une modification Y_n^2 de la Nikulin-Rao-Robson (NRR) statistique proposée par Bagdonaviciuc et Nikulin (2011). Elle est basée sur les estimateurs de maximum de vraisemblance sur les données non regroupées et tend à la limite vers une distribution du Chi-deux. Nous donnons la forme explicite de tous les éléments des matrices W, I, G , qui composent la statistique de test.

Dans le dernier chapitre, nous présentons une étude détaillée par simulations numériques pour confirmer les résultats obtenus. Nous avons simulé 10.000 échantillons de différentes tailles et avec différentes valeurs de paramètres. Nous avons calculé les estimateurs du maximum de vraisemblance et la statistique de test correspondante pour chacun de ces échantillons. En-

fin, nous terminons par des exemples issus de données réelles de fiabilité et de données de survie.

Chapitre 1

Présentation du modèle inverse Weibull généralisé (IWG)

la distribution inverse Weibull généralisée (IWG) est une loi de probabilité récente, proposée par Gusmao et al. (2009), généralisant la distribution connue et très utilisée de l'inverse Weibull. Ayant trois paramètres, elle est plus flexible que cette dernière, et grâce à ses interprétations physiques, elle peut être appliquée dans différents domaines tels que la fiabilité, l'analyse de survie, les sciences des matériaux, la psychologie,.... .

1.1 Quelques définitions

1.1.1 Censure

La censure est un phénomène que l'on rencontre lors du recueil des données de survie. Pour un individu i , on considère :

- son temps de survie X_i ,
- son temps de censure C_i ,
- la durée réellement observée T_i .

Censure à droite

La durée de vie est dite censurée à droite si l'individu n'a pas subi l'événement à sa dernière observation.

Censure de type I

Étant donné un échantillon $X = (X_1, \dots, X_n)^T$ de durées de vie X et un nombre positif fixe C . On dit qu'il y a une censure à droite de cet échantillon si au lieu d'observer X_1, \dots, X_n , on observe X_i , uniquement si $X_i \leq C$, sinon on sait seulement que $X_i > C$. On note alors $T_i = \min(X_i, C)$.

On rencontre ce type de censure en industrie.

Censure de type II

Étant donné un échantillon $X = (X_1, \dots, X_n)^T$ de durées de vie X_i et un nombre entier positif r , on dit qu'il y a censure de type II si au lieu d'observer X_1, \dots, X_n , on observe n statistiques

$$(T_1, \delta_1), \dots, (T_n, \delta_n), \quad T_i = X_i \wedge X_{(r)}, \quad \delta_i = 1_{\{T_i=X_i\}}$$

La date de censure est $X_{(r)}$ est la r -ième statistique d'ordre.

Censure de type III ou censure aléatoire

Étant donné un échantillon $X = (X_1, \dots, X_n)^T$ de durées de vie X_i , on dit qu'il y a censure aléatoire de cet échantillon s'il existe un autre échantillon $C = (C_1, \dots, C_n)^T$ indépendant de X tel qu'au lieu d'observer X_1, \dots, X_n , on observe n statistiques $(T_1, \delta_1), \dots, (T_n, \delta_n)$ avec :

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

Donc en cas de censure aléatoire, on associe à chaque sujet i une statistique (X_i, C_i) de dimension 2, dont seulement la plus petite composante est observée.

Définition de la censure à gauche

La censure à gauche correspond au cas où l'individu a déjà subi l'évènement avant qu'on ne l'observe. On sait seulement que la date de l'évènement T est inférieure à un certain délai connu C .

On associe à chaque individu un couple de variables aléatoires (X, δ) :

$$X = \max(T, C), \quad \delta = 1_{T < C}$$

1.1.2 Fonction de vraisemblance

Soit $X = (X_1, X_2, \dots, X_n)^T$ un n -échantillon de variables aléatoires i.i.d de loi $f(x, \theta)$ dont le paramètre θ est inconnu, on appelle vraisemblance de X la fonction L_n définie par :

$$L_n(\theta) = \prod_{i=1}^n f(x_i, \theta), \quad \theta \in \Theta.$$

Dans le cas d'une censure à droite, C indépendante de la durée d'intérêt X , la distribution de X est définie par un paramètre de dimension finie ; toute l'information est contenue dans le couple (T_i, δ_i) , où $T_i = \min(X_i, C_i)$ est la durée observée, et l'indicateur de censure $\delta_i = 1_{X_i \leq C_i}$.

Par l'hypothèse de censure non informative, le paramètre d'intérêt θ n'apparaît pas dans la loi de la censure et la contribution à la vraisemblance de l'individu i se réduit à :

$$L(\theta) = \prod_{i=1}^n f(t_i, \theta)^{\delta_i} S(t_i, \theta)^{1-\delta_i}$$

1.1.3 Estimateur de Kaplan-Meier

L'estimateur de *Kaplan-Meier* de la fonction de survie à partir d'un échantillon (T_i, D_i) , $i = 1, \dots, n$ d'observations éventuellement censurées est donné par :

$$\hat{S}_{KM}(t) = \prod_{i=1}^n \left(1 - \frac{I(T_i \leq t, D_i = 1)}{Y(T_i)} \right)$$

où $Y(t)$ représente le nombre de sujets à risque (ni censurés, ni décédés) à l'instant t et où $I(A)$ représente la fonction indicatrice de l'événement A

et vaut 1 si l'évènement A est réalisé et 0 sinon. Une autre écriture de l'estimateur de Kaplan-Meier est la suivante. Ordonnons les durées observées non-censurées : $T'_1 < \dots < T'_k$ dans l'intervalle $[T'_j, T'_{j+1}[$. Soit n_j le nombre d'individus à risque (ni morts, ni censurés) à l'instant T'_j .

L'estimateur de Kaplan-Meier se définit alors comme :

$$\begin{aligned}\hat{S}_{KM}(t) &= \prod_{j=1}^k \left(\frac{n_j - d_j}{n_j} \right) \quad \text{où } k \text{ est défini par } T'_k \leq t < T'_{k+1} \\ \hat{S}_{KM}(t) &= 1 \quad \text{pour } t < T'_1\end{aligned}$$

où

n_j est le nombre de survivants juste avant le temps t_j , lorsqu'il n'y aucune censure.

n_j est le nombre de survivants moins le nombre de pertes, lorsqu'il y a censure.

d_j est le nombre de décès au temps t_j .

Les intervalles de confiance pour $\Lambda(t)$ permettent d'obtenir ceux pour $S(t)$.

Les intervalles de confiance pour $\Lambda(t)$ sont donnés par :

$$\Lambda(t) \in \left[\widehat{\Lambda}(t) \pm z_{1-\alpha/2} \cdot \widehat{se}(\widehat{\Lambda}(t)) \right]$$

Les intervalles de confiance pour la survie $S(t)$ sont obtenus par exponentiation des intervalles de confiance pour $\Lambda(t)$ sur la base de la formule

$S(t) = \exp(-\Lambda(t))$:

$$S(t) \in \left[\exp \left(-\widehat{\Lambda}(t) \pm z_{1-\alpha/2} \cdot \widehat{se}(\widehat{\Lambda}(t)) \right) \right]$$

où $\widehat{\Lambda}$ représente l'estimateur de Nelson-Aalen de Λ :

$$\widehat{\Lambda}(t) = \sum_{i=1}^n \left(\frac{I(T_i \leq t, D_i = 1)}{Y(t_i)} \right)$$

que l'on peut aussi écrire sous la forme suivante :

$$\begin{aligned}\widehat{\Lambda}(t) &= \sum_{j=1}^k \frac{d_j}{n_j} \quad \text{où } k \text{ est défini par } T'_k \leq t < T'_{k+1} \\ \widehat{\Lambda}(t) &= 0 \quad \text{pour } t < T'_1\end{aligned}$$

L'estimateur de Tsiatis (ou d'Aalen) de la variance de $\widehat{\Lambda}$ est donné par :

$$\begin{aligned}\widehat{Var}(\widehat{\Lambda}(t)) &= \sum_{j=1}^k \frac{d_j}{n_j^2} \quad \text{où } k \text{ est défini par } T'_k \leq t < T'_{k+1} \\ \widehat{Var}(\widehat{\Lambda}(t)) &= 0 \quad \text{pour } t < T'_1\end{aligned}$$

L'estimateur de Greenwood de la variance de $\widehat{\Lambda}$ est donné par :

$$\begin{aligned}\widehat{Var}(\widehat{\Lambda}(t)) &= \sum_{j=1}^k \frac{d_j}{n_j(n_j - d_j)} \quad \text{où } k \text{ est défini par } T'_k \leq t < T'_{k+1} \\ \widehat{Var}(\widehat{\Lambda}(t)) &= 0 \quad \text{pour } t < T'_1\end{aligned}$$

1.2 La distribution inverse Weibull (IW)

Keller et Kamath (1982) étudient les formes de la densité et les fonctions du taux de défaillance pour le modèle de base inverse Weibull. Plus tard, Drapella (1993), Mudholkar et Kollia (1994) l'ont nommée distribution complémentaire de Weibull ou distribution réciproque de Weibull.

La fonction de distribution cumulative de l'IW prend la forme :

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

où $\alpha > 0$ et $\beta > 0$.

La fonction densité de probabilité correspondante est :

$$g(t) = \beta\alpha^\beta t^{-(\beta+1)} \exp\left[-\left(\frac{\alpha}{t}\right)^\beta\right]$$

la fonction de survie de l'IW est donnée par :

$$S(t) = 1 - G(t) = 1 - \exp\left[-\left(\frac{\alpha}{t}\right)^\beta\right]$$

Le taux de hasard pour l'IW est :

$$\lambda(t) = \beta\alpha^\beta t^{-(\beta+1)} \exp\left[-\left(\frac{\alpha}{t}\right)^\beta\right] \left\{1 - \exp\left[-\left(\frac{\alpha}{t}\right)^\beta\right]\right\}^{-1}$$

Le $k^{ième}$ moment s'écrit :

$$E(T^k) = \alpha^k \Gamma(1 - k\beta^{-1}),$$

où $\Gamma(\cdot)$ est la fonction gamma.

Pour simuler la distribution IW , nous pouvons utiliser l'équation non linéaire :

$$t = \alpha \left[\frac{-\log(u)}{\gamma} \right]^{-1/\beta},$$

où u suit la distribution uniforme $U(0, 1)$.

Quelques exemples de la densité IW sont présentés respectivement dans les figures 1 et 2, et le taux de hasard dans les figures 3, 4 et 5.

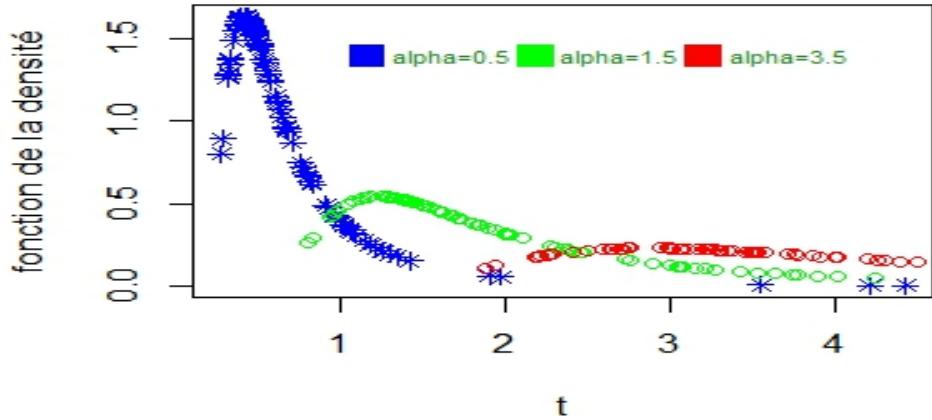


Fig. 1. Densité de l'IW pour le paramètre de forme $\beta = 2$ avec $\alpha = (0.5, 1.5, 3.5)$.

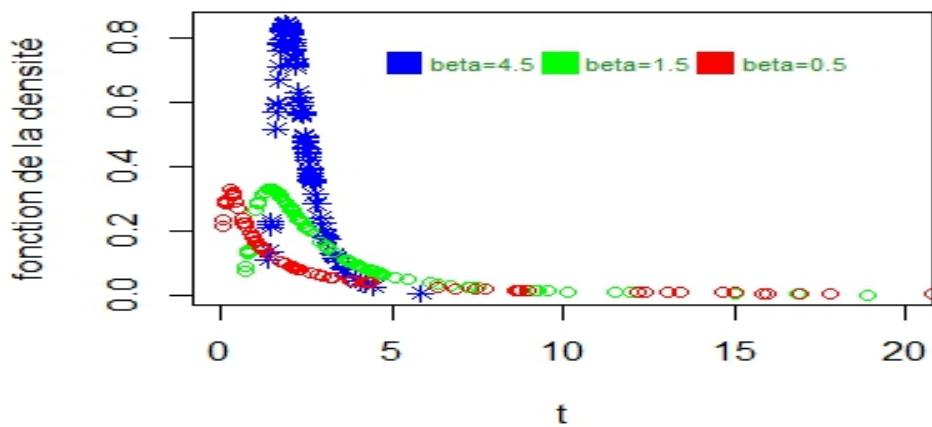


Fig. 2. Densité de l'IW pour le paramètre de position $\alpha = 2$ avec $\beta = (0.5, 1.5, 4.5)$.

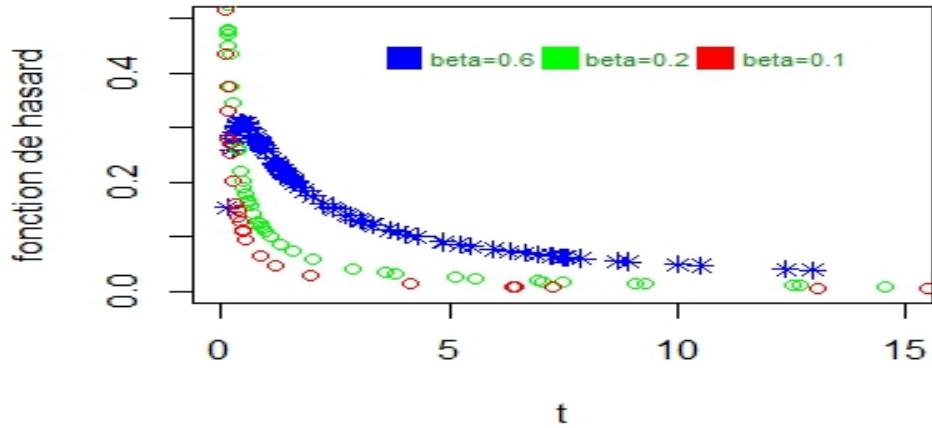


Fig. 3. Taux de hasard de l'IW pour le paramètre de position $\alpha = 2$ avec $\beta = (0.6, 0.2, 0.1)$.

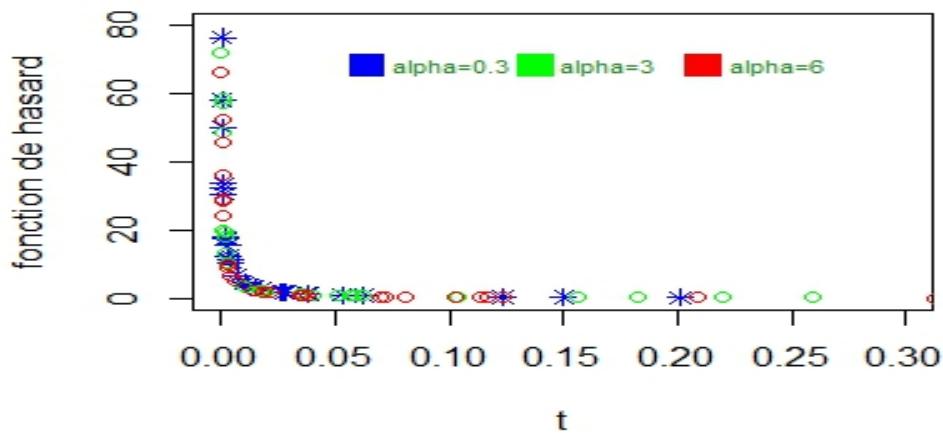


Fig. 4. Taux de hasard de l'IW pour le paramètre de forme $\beta = 0.1$ avec $\alpha = (0.3, 3, 6)$.

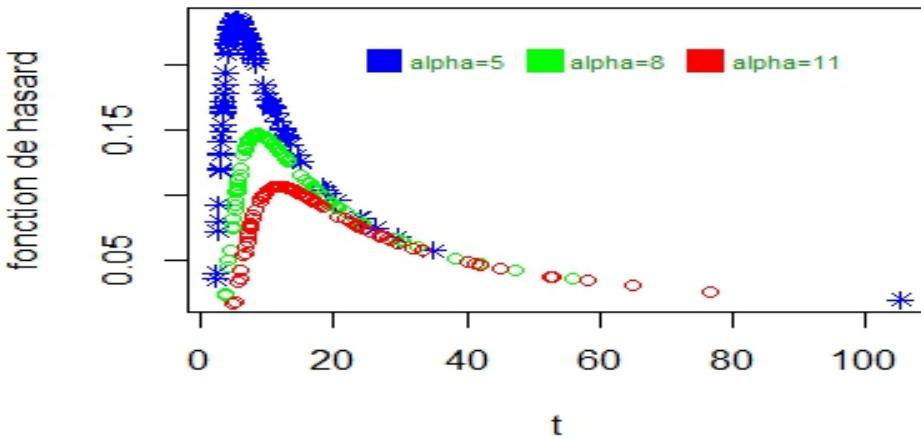


Fig. 5. Taux de hasard de l'IW pour le paramètre de forme $\beta = 2$ avec $\alpha = (5, 8, 11)$.

Nous remarquons que :

- Si $\beta \in [0, 0.3[$, le taux de panne est décroissant.
- Si $\beta \in [0.3, 2[$, le taux de panne de l'IW est croissant.
- Si $\beta \geq 2$, le taux de panne de l'IW est en forme de cloche (cup shape).

1.3 La distribution inverse Weibull généralisée (IWG)

La fonction de répartition de la distribution inverse Weibull généralisée (IWG) a été introduite par Felipe R. S. de Gusmao, Edwin M. M. Ortega et Gauss M. Cordeiro (2009) en élevant la fonction de répartition de la distribution de l'IW notée par $G(t)$ à la puissance $\gamma > 0$, comme suit :

$$F(t) = G(t)^\gamma = \exp \left[-\gamma \left(\frac{\alpha}{t} \right)^\beta \right] \quad (1.1)$$

Ainsi, la fonction densité de l'IWG dépend de trois paramètres $\alpha > 0$, $\beta > 0$ et $\gamma > 0$ et est donnée par :

$$f(t) = \gamma\beta\alpha^\beta t^{-(\beta+1)} \exp\left[-\gamma\left(\frac{\alpha}{t}\right)^\beta\right], \quad t > 0 \quad (1.2).$$

Si $\gamma = 1$, alors la distribution de l'IW est un cas particulier de l'IWG.

La fonction de survie et le taux de hasard sont respectivement :

$$S(t) = 1 - F(t) = 1 - \exp\left[-\gamma\left(\frac{\alpha}{t}\right)^\beta\right] \quad (1.3)$$

et

$$\lambda(t) = \gamma\beta\alpha^\beta t^{-(\beta+1)} \exp\left[-\gamma\left(\frac{\alpha}{t}\right)^\beta\right] \left\{1 - \exp\left[-\gamma\left(\frac{\alpha}{t}\right)^\beta\right]\right\}^{-1}$$

Pour simuler la distribution IWG, on utilise l'équation non linéaire

$$t = \alpha \left[\frac{-\log(u)}{\gamma} \right]^{-1/\beta},$$

où u est de distribution uniforme $U(0, 1)$.

Des graphes de la densité de l'IWG sont présentés respectivement dans les figures 6, 7 et 8, et le taux de hasard dans les figures 9, 10 et 11.

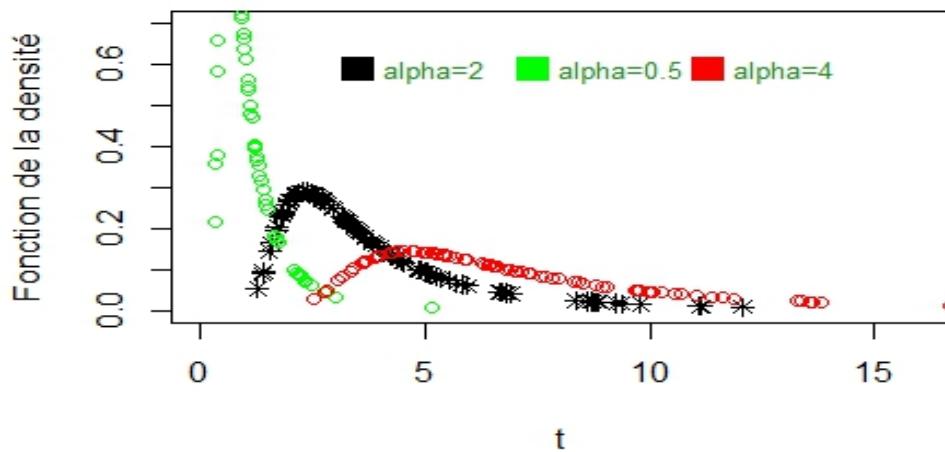


Fig. 6. Densité de l'IWG pour $\beta = 2$ et $\gamma = 2$ avec $\alpha = (0.5, 2, 4)$.

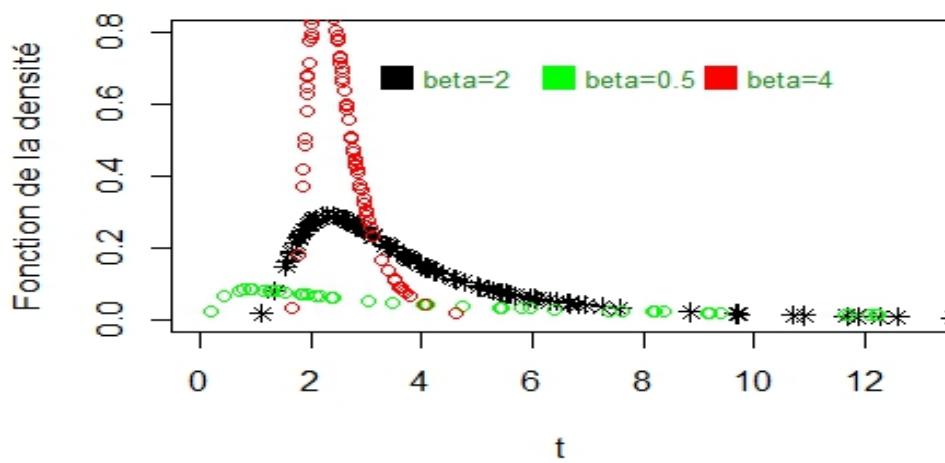


Fig. 7. Densité de l'IWG pour $\alpha = 2$ et $\gamma = 2$ avec $\beta = (0.5, 2, 4)$.

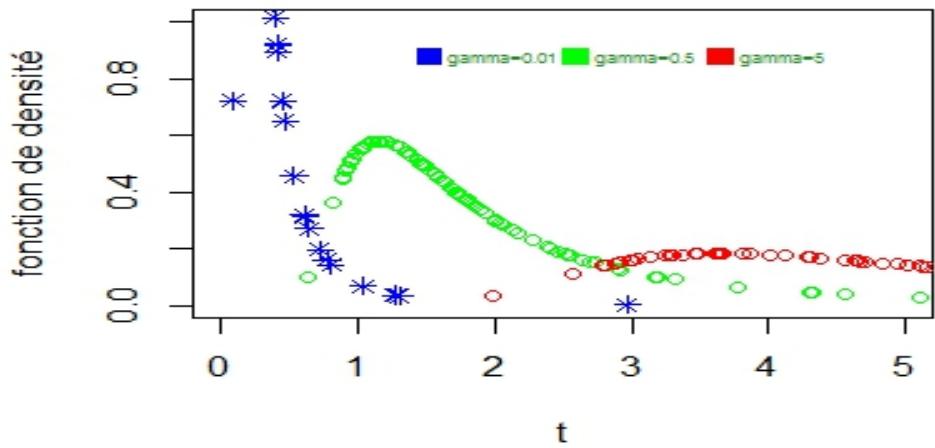


Fig. 8. Densité de l'IWG pour $\alpha = 2$ et $\beta = 2$ avec $\gamma = (0.01, 0.5, 5)$.

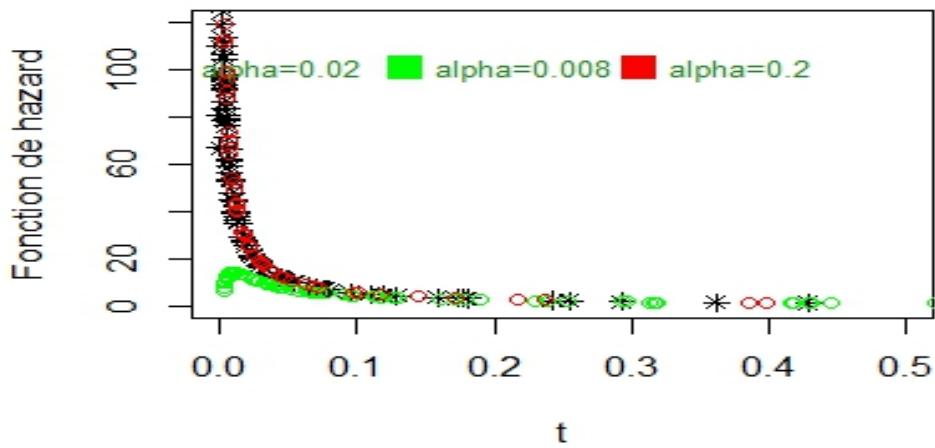


Fig. 9. Taux de hasard de l'IWG pour $\beta = 2$ et $\gamma = 2$ avec $\alpha = (0.02, 0.008, 0.2)$.

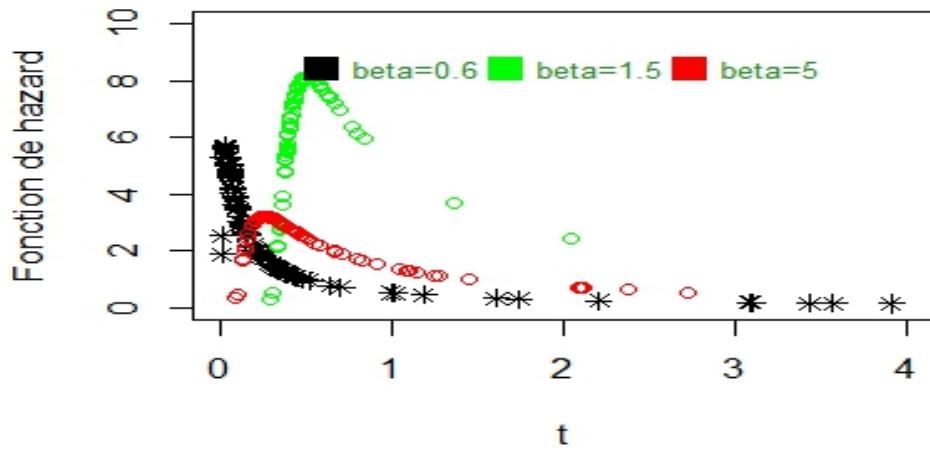


Fig. 10. Taux de hasard de l'IWG pour $\alpha = 2$ et $\gamma = 2$ avec $\beta = (0.62, 1.5, 5)$.

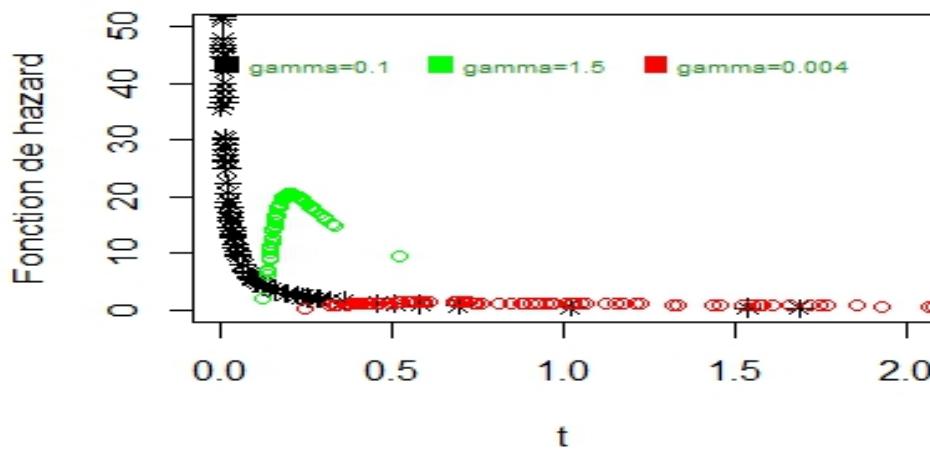


Fig. 11. Taux de hasard de l'IWG pour $\alpha = 2$ et $\beta = 2$ avec $\gamma = (0.004, 0.1, 1.5)$.

Nous remarquons que la forme du taux de hasard de la distribution IWG est en cloche (unimodale) et monotone en fonction des valeurs des paramètres. C'est une distribution très flexible et peut décrire différents taux de panne en fiabilité ou en analyse de survie.

1.4 Fonction génératrice des moments

Soit la variable aléatoire T , de densité IWG (1.2), le $K^{ième}$ moment est donné par :

$$E(T^k) = \gamma^{\frac{k}{\beta}} \alpha^k \Gamma(1 - k\beta^{-1})$$

La fonction génératrice des moments $M(z)$ de T pour $|z| < 1$ est :

$$M(z) = E(e^{zT}) = \sum_{k=0}^n \left\{ \frac{z^k}{k!} \gamma^{\frac{k}{\beta}} \alpha^k \Gamma(1 - k\beta^{-1}) \right\}$$

Ainsi, pour $|z| < 1$, la fonction génératrice des cumulants de T est :

$$K(z) = \log [M(z)] = \log \left\{ \sum_{k=0}^n \left[\frac{z^k}{k!} \gamma^{\frac{k}{\beta}} \alpha^k \Gamma(1 - k\beta^{-1}) \right] \right\}.$$

Les quatre premiers cumulants sont :

$$k_1 = \gamma^{\frac{1}{\beta}} \alpha \Gamma_1(\beta), \quad k_2 = \gamma^{\frac{2}{\beta}} \alpha^2 \{ \Gamma_2(\beta) - \Gamma_1(\beta)^2 \},$$

$$k_3 = \gamma^{\frac{3}{\beta}} \alpha^3 \{ \Gamma_3(\beta) - 3\Gamma_2(\beta) \Gamma_1(\beta) + 2\Gamma_1(\beta)^3 \}$$

et

$$k_4 = \gamma^{\frac{4}{\beta}} \alpha^4 \left\{ \Gamma_4(\beta) - 4\Gamma_3(\beta)\Gamma_1(\beta) - 3\Gamma_1(\beta)^2 + 12\Gamma_2(\beta)\Gamma_1(\beta)^2 - 6\Gamma_1(\beta)^4 \right\},$$

où $\Gamma_k(\beta) = \Gamma(1 - k\beta^{-1})$ pour $k = 1, \dots, 4$.

Les coefficients d'asymétrie et d'aplatissement sont respectivement :

$$\rho_3 = k_3/k_2^{\frac{3}{2}}$$

et

$$\rho_4 = k_4/k_2^2$$

1.5 Estimation du maximum de vraisemblance en cas de données complètes

1.5.1 Estimation du maximum de vraisemblance

Si une variable aléatoire $t \sim IWG(\theta)$ de vecteur de paramètres $\theta = (\alpha, \beta, \gamma)$, alors la fonction densité est définie comme suit :

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

Afin de calculer les estimateurs du maximum de vraisemblance des paramètres inconnus de cette distribution, on a besoin de la fonction de vraisemblance donnée par l'équation suivante :

$$\begin{aligned}
L(t, \theta) &= \prod_{i=1}^n f(t_i, \theta) \\
&= \prod_{i=1}^n \gamma \beta \alpha^\beta t_i^{-(\beta+1)} \exp \left[-\gamma \left(\frac{\alpha}{t_i} \right)^\beta \right] \\
&= \gamma^n \beta^n \alpha^{n\beta} \prod_{i=1}^n t_i^{-(\beta+1)} \exp \left[-\gamma \left(\frac{\alpha}{t_i} \right)^\beta \right]
\end{aligned}$$

Et l'expression de la fonction log-vraisemblance est :

$$\begin{aligned}
\ell(t, \theta) &= \sum_{i=1}^n \ln f(t_i, \theta) \\
&= n \ln \beta + n \ln \gamma + n \beta \ln \alpha - (\beta + 1) \sum_{i=1}^n \ln t_i - \sum_{i=1}^n \gamma \left(\frac{\alpha}{t_i} \right)^\beta
\end{aligned}$$

Nous calculons les fonctions de score :

$$\frac{\partial}{\partial \theta} \ell(\theta) = \left(\frac{\partial}{\partial \alpha} \ell(\theta), \frac{\partial}{\partial \beta} \ell(\theta), \frac{\partial}{\partial \gamma} \ell(\theta) \right)^T$$

Pour notre distribution IWG, les fonctions de score sont obtenues comme suit :

$$\begin{aligned}
\frac{\partial \ell(\theta)}{\partial \alpha} &= \frac{n\beta}{\alpha} - \gamma \beta \alpha^{\beta-1} \sum_{i=1}^n t_i^{-\beta} = 0 \\
\frac{\partial \ell(\theta)}{\partial \beta} &= \frac{n}{\beta} + n \ln \alpha - \sum_{i=1}^n \ln t_i - \gamma \alpha^\beta \sum_{i=1}^n t_i^{-\beta} \ln \left(\frac{\alpha}{t_i} \right) = 0 \\
\frac{\partial \ell(\theta)}{\partial \gamma} &= \frac{n}{\gamma} - \alpha^\beta \sum_{i=1}^n t_i^{-\beta} = 0
\end{aligned}$$

Puisque, nous ne pouvons pas trouver les formules explicites des estimateurs du maximum de vraisemblance des paramètres, nous utilisons alors des méthodes numériques telles que la méthode de Newton Raphson, la méthode de Monte Carlo, l'algorithme BB ou autres.

1.5.2 Matrice d'information de Fisher

Soit $T = (t_1, \dots, t_n)$ un échantillon aléatoire, et soit $f(t, \theta)$ la fonction densité de probabilité pour un certain modèle de données, de vecteur de paramètres $\theta = (\theta_1, \dots, \theta_k)$, alors la matrice d'information de Fisher $I_n(\theta)$ sur l'échantillon de taille n est donnée par la matrice symétrique $k \times k$.

$$I_n(\theta)_{i,j} = Cov \left[\frac{\partial \ln f(t, \theta)}{\partial \theta_i}, \frac{\partial f(t, \theta)}{\partial \theta_j} \right]$$

ou bien

$$I_n(\theta)_{i,j} = -E \left[\frac{\partial^2 \ln f(t, \theta)}{\partial \theta_i \partial \theta_j} \right]$$

Après de longs calculs, nous obtenons les éléments de la matrice d'information de Fisher de la distribution inverse Weibull généralisée IWG :

$$\frac{\partial^2 \ell(\theta)}{\partial \alpha^2} = \frac{-n\beta}{\alpha^2} - \beta\gamma(\beta-1)\alpha^{\beta-2} \sum_{i=1}^n t_i^{-\beta}$$

$$\frac{\partial^2 \ell(\theta)}{\partial \alpha \partial \beta} = \frac{n}{\alpha} - \gamma\alpha^{\beta-1} \sum_{i=1}^n t_i^{-\beta} - \gamma\beta\alpha^{\beta-1} \ln \alpha \sum_{i=1}^n t_i^{-\beta} + \gamma\beta\alpha^{\beta-1} \sum_{i=1}^n t_i^{-\beta} \ln t_i$$

$$\frac{\partial^2 \ell(\theta)}{\partial \alpha d\gamma} = -\beta\alpha^{\beta-1} \sum_{i=1}^n t_i^{-\beta}$$

$$\begin{aligned}\frac{\partial^2 \ell(\theta)}{\partial \beta^2} &= -\frac{n}{\beta^2} - \gamma \sum_{i=1}^n \ln\left(\frac{\alpha}{t_i}\right) \ln\left(\frac{\alpha}{t_i}\right) \left(\frac{\alpha}{t_i}\right)^\beta \\ \frac{\partial^2 \ell(\theta)}{\partial \beta \partial \gamma} &= - \sum_{i=1}^n \left(\frac{\alpha}{t_i}\right)^\beta \ln\left(\frac{\alpha}{t_i}\right) \\ \frac{\partial^2 \ell(\theta)}{\partial \gamma^2} &= -\frac{n}{\gamma^2}\end{aligned}$$

1.6 Estimation du maximum de vraisemblance en cas de données censurées

Soit t une variable aléatoire distribuée selon une loi IWG de vecteur de paramètres $\theta = (\alpha, \beta, \gamma)^T$. Chaque individu i est supposé avoir une durée de vie T_i et un temps de censure C_i , où T_i et C_i sont des variables aléatoires indépendantes. Supposons que les données sont constituées de n observations indépendantes $t_i = \min(T_i, C_i)$ pour $i = 1, \dots, n$. La distribution de C_i ne dépend pas des paramètres inconnus de T_i .

la fonction de vraisemblance dans le cas de données censurées peut être donnée par :

$$L(t, \theta) = \prod_{i=1}^n f^{\delta_i}(t_i, \theta) S^{1-\delta_i}(t_i, \theta), \quad \delta_i = 1_{\{T_i \leq C_i\}}$$

La fonction log-vraisemblance est :

$$\ell(t, \theta) = \sum_{i=1}^n \delta_i \ln f(t_i, \theta) + \sum_{i=1}^n (1 - \delta_i) \ln S(t_i, \theta)$$

Comme

$$\lambda(t, \theta) = \frac{f(t, \theta)}{S(t, \theta)}$$

alors l'équation de la log-vraisemblance devient :

$$\ell(\theta) = \sum_{i=1}^n \delta_i \ln \lambda(t_i, \theta) + \sum_{i=1}^n \ln S(t_i, \theta)$$

Dans notre cas, l'équation de la log-vraisemblance devient :

$$\begin{aligned} \ell(t, \theta) &= \sum_{i=1}^n \delta_i \ln \left(\beta \alpha^\beta t_i^{-(\beta+1)} \exp \left[-\left(\frac{\alpha}{t_i} \right)^\beta \right] \left\{ 1 - \exp \left[-\left(\frac{\alpha}{t_i} \right)^\beta \right] \right\}^{-1} \right) \\ &\quad + \sum_{i=1}^n \ln \left(1 - \exp \left[-\gamma \left(\frac{\alpha}{t_i} \right)^\beta \right] \right) \end{aligned}$$

Après simplifications, nous obtenons :

$$\begin{aligned} \ell(t, \theta) &= r[\log(\gamma) + \log(\beta) + \beta \log(\alpha)] - (\beta + 1) \sum_{i \in F} \log(t_i) - \gamma \alpha^\beta \sum_{i \in F} t_i^{-\beta} \\ &\quad + \sum_{i \in C} \log \left\{ 1 - \exp \left[-\gamma \left(\frac{\alpha}{t_i} \right)^\beta \right] \right\} \end{aligned} \quad (1.4)$$

où r est le nombre de défaillances, F et C désignent respectivement les ensembles d'observations non censurées et censurées.

Les fonctions de score pour les paramètres α , β et γ sont données par :

$$\frac{\partial \ell(\theta)}{\partial \alpha} = \frac{r\beta}{\alpha} - \gamma\beta\alpha^{\beta-1} \sum_{i \in F} t_i^{-\beta} + \gamma\beta\alpha^{\beta-1} \sum_{i \in C} t_i^{-\beta} \left(\frac{1 - \left[1 - \exp \left[-\gamma \left(\frac{\alpha}{t_i} \right)^\beta \right] \right]}{\left[1 - \exp \left[-\gamma \left(\frac{\alpha}{t_i} \right)^\beta \right] \right]} \right)$$

$$\begin{aligned} \frac{\partial \ell(\theta)}{\partial \beta} &= \frac{r}{\beta} + r \log(\alpha) - \sum_{i \in F} \log(t_i) - \gamma\alpha^\beta \sum_{i \in F} t_i^{-\beta} \log \left(\frac{\alpha}{t_i} \right) \\ &\quad + \gamma\alpha^\beta \sum_{i \in C} t_i^{-\beta} \log \left(\frac{\alpha}{t_i} \right) \left(\frac{1 - \left[1 - \exp \left[-\gamma \left(\frac{\alpha}{t_i} \right)^\beta \right] \right]}{\left[1 - \exp \left[-\gamma \left(\frac{\alpha}{t_i} \right)^\beta \right] \right]} \right), \end{aligned}$$

$$\frac{\partial \ell(\theta)}{\partial \gamma} = \frac{r}{\gamma} - \alpha^\beta \sum_{i \in F} t_i^{-\beta} + \alpha^\beta \sum_{i \in C} t_i^{-\beta} \left(\frac{1 - \left[1 - \exp \left[-\gamma \left(\frac{\alpha}{t_i} \right)^\beta \right] \right]}{\left[1 - \exp \left[-\gamma \left(\frac{\alpha}{t_i} \right)^\beta \right] \right]} \right).$$

Pour résoudre le système de fonctions de score, assez compliquées en présence de la censure, on fait appel aux méthodes numériques, telles que la méthode de Monte Carlo, l'algorithme BB ou autres.

Chapitre 2

Construction du modèle AFT-IWG

L’analyse des données de survie est l’étude du délai de la survenue d’un événement précis comme par exemple l’apparition d’une maladie en analyse de survie, ou bien la panne d’une machine (durée de fonctionnement d’une machine) en fiabilité, ou encore la survenue d’un sinistre en actuariat. Les statistiques développent des techniques et des méthodes qui permettent d’analyser les données issues de l’observation, afin de cerner les caractéristiques de la population concernée et d’identifier un modèle capable d’engendrer ces données.

Cox en 1972, a proposé un modèle simple et n’a fait aucune hypothèse sur les fonctions de risque de base, il a seulement supposé que ces fonctions de risque sont proportionnelles. C’est pourquoi on l’appelle aussi ”proportional hazard model” (**PH**). Cependant ceci n’est pas toujours réalisable ; en plus l’effet des variables explicatives est mesuré sur la fonction risque. Par contre dans les modèles de durées de vie accélérées AFT, on mesure cet effet sur le temps de survie. Ceci permet de mieux interpréter les résultats et de considérer que le modèle AFT est une bonne alternative au modèle PH.

Dans ce cas et pour estimer la fonction de fiabilité sous les stress usuels, il devient nécessaire d’effectuer des expériences accélérées (AFT) qui permettent de raccourcir la durée de vie du système en lui appliquant des stress élevés afin d’observer les pannes dans un temps raisonnable.

Les modèles AFT peuvent être paramétriques, semi-paramétriques ou non-paramétriques selon que la fonction de survie de base appartienne ou non à une famille paramétrique et que la fonction de stress soit inconnue ou non. Ces modèles sont utilisés aujourd’hui non seulement en fiabilité et en analyse de survie, mais aussi en dynamique des populations, biologie, génétique, radio-biologie, biophysique, vieillissement, dégradation, etc ...

Plusieurs modèles paramétriques AFT (AFT-exponentiel, AFT-Weibull, AFT-lognormal, AFT-inverse gaussien) ont été étudiés (voir Bagdonavičius et Nikulin (2002)).

Dans ce chapitre, nous nous intéressons à la construction d’un modèle AFT dont la distribution de base est une distribution inverse Weibull généralisée. Ce modèle est très intéressant car il peut décrire différentes durées de vie dans différents domaines. Nous estimons les paramètres, nous donnons tous les éléments de la matrice d’information de Fisher, et nous calculons l’intervalle de confiance de la fonction de survie du modèle AFT.

2.1 Introduction des modèles AFT

Le principal objectif des modèles AFT est de connaître l’influence des covariables sur la durée de vie des individus. Le modèle AFT permet de contrôler le processus de dégradation en mettant les unités sous stress, afin d’estimer les différentes caractéristiques du modèle et dégager de meilleurs estimateurs ; et de ce fait améliorer la qualité des produits et optimiser leurs maintenances pour pouvoir affronter la concurrence particulièrement en industrie.

les types de stress sont variés selon le domaine d’étude. Nous en citons quelques uns.

- **Les stress constants** : les stress restent constants au cours du temps, chaque élément est testé à un niveau de stress constant.

- **Les stress progressifs** : ces types de stress augmentent de façon continue avec le temps.

- **Les stress variables** : les produits sont testés sous des stress variés est continus avec le temps.

Les stress variant dans le temps les plus fréquemment utilisés sont **les stress en escalier** : les unités sont mises au banc d'essai à un stress initial faible et si elles ne tombent pas en panne dans un temps prédéterminé t_1 , le stress est augmenté. Si les unités ne tombent pas encore en panne dans un temps prédéterminé $t_2 > t_1$, le stress est augmenté à nouveau, et ainsi de suite.

Notons par E_m le sous-ensemble de E des stress en escalier à m niveaux de la forme :

$$z(t) = \begin{cases} z_1, & 0 \leq t < t_1 \\ z_2, & t_1 \leq t < t_2 \\ . & . \\ z_m, & t_{m-1} \leq t < t_m = +\infty \end{cases} \quad (2.1)$$

où z_1, z_2, \dots, z_m appartiennent à E .

Le sous-ensemble E_2 de E prend la forme :

$$z(t) = \begin{cases} z_1, & 0 \leq t < t_1 \\ z_2, & t \geq t_1 \end{cases} \quad (2.2)$$

Considérons E l'ensemble de tous les stress possibles, défini par :

$$E = \left\{ z(\cdot) = (z_1(\cdot), z_2(\cdot), \dots, z_m(\cdot))^T \right\}$$

Si $z(\cdot)$ est constant, nous écrivons z à la place de $z(\cdot)$, et nous notons par E_1 le sous-ensemble de E de tous les stress constants.

Notons par $T_{z(\cdot)}$ le taux de panne sous $z(\cdot)$ et par $f_{z(\cdot)}(t)$, $S_{z(\cdot)}(t)$ et $F_{z(\cdot)}(t)$ respectivement la fonction de densité, la fonction de survie et la fonction cumulative de la distribution, où :

$$S_{z(\cdot)}(t) = P(T_{z(\cdot)} \geq t) = 1 - F_{z(\cdot)}(t), \quad z(\cdot) \in E$$

La fonction risque de $T_{z(.)}$ sous $z(.)$ est :

$$\lambda_{z(.)}(t) = \lim_{h \rightarrow 0} \frac{1}{h} P(t \leq T_{z(.)} < t + h \setminus T_{z(.)} \geq t) = -\frac{S'_{z(.)}(t)}{S_{z(.)}(t)}$$

Notons par

$$\Lambda_{z(.)}(t) = \int_0^t \lambda_{z(.)}(u) du = -\ln(S_{z(.)}(t))$$

la fonction cumulative de risque de $T_{z(.)}$.

2.1.1 Définition du modèle AFT

Définition :

Le modèle AFT est défini sur E , si la fonction de survie sous stress $z(.) \in E$, est

$$S_{z(.)}(t) = S_0 \left(\int_0^t r(z(u)) du \right), \forall z(.) \in E \quad (2.3)$$

Si $z(.) = z$ est constant, alors le modèle (2.3) devient :

$$S_z(t) = S_0(r(z)t), \forall z(.) \in E_1$$

et sur E_m , l'équation (2.3) s'écrit [Bagdonavičius et Nikulin (2002)] :

$$\begin{aligned} S_z(t) &= S_0 \left\{ \sum_{j=1}^{i-1} r(z_j) (t_j - t_{j-1}) + r(z_i) (t - t_{i-1}) \right\} \\ &= S_{z_i} \left\{ t - t_{i-1} + \frac{1}{r(z_i)} \sum_{j=1}^{i-1} r(z_j) (t_j - t_{j-1}) \right\}, t \in [t_{i-1}, t_i[, i = 1, \dots, m. \end{aligned}$$

La fonction $r(z)$ peut être choisie parmi certaines classes de fonctions.
Souvent, le modèle AFT est paramétrisé comme :

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

où $\beta = (\beta_0, \dots, \beta_m)^T$ est le vecteur de régression inconnu. Et

$$z = (z_0, \dots, z_m) = (\Psi_0(z), \dots, \Psi_m(z))^T$$

est le vecteur des fonctions de stress inconnues Ψ_i où $\Psi_0(t) = 1$.

Ψ est une fonction de z , cette fonction peut prendre plusieurs formes :

$r(z) = \exp \{-\beta_0 - \beta_1 z\}$, $\Psi(z) = z$, c'est le modèle log-linéaire.

$r(z) = \exp \{-\beta_0 - \beta_1/z\}$, $\Psi(z) = 1/z$, c'est le modèle d'Arrhenius.

$r(z) = \exp \left\{ -\beta_0 - \beta_1 \ln \frac{z}{1-z} \right\}$, $0 < z < 1$, $\Psi(z) = \ln \frac{z}{1-z}$, modèle de Meeker-Luvalle.

Le modèle AFT paramétrique est donné par la formule :

$$S_{z(\cdot)}(t) = S_0 \left(\int_0^t \exp \{-\beta^T z(u)\} du \right), \quad z(\cdot) \in E \quad (2.4)$$

Si les stress sont constants dans le temps, on a :

$$S_{z(\cdot)}(t) = S_0 \left(\exp \{-\beta^T z\} t \right)$$

Le logarithme du taux de défaillance T_z sous z peut s'écrire :

$$\ln(T_z) = \beta^T z + \epsilon$$

où la fonction de survie de ϵ ne dépend pas de z et est $S(t) = S_0(\ln t)$.

2.1.2 Plans d'expérience

Comme on l'a mentionné, le but des tests de vie accélérés (ALT) est de donner l'estimateur de la fonction de fiabilité $S_{z^{(0)}}$ sous le stress usuel $z^{(0)}$, en utilisant les données d'expérience accélérées lorsque les unités sont testées sous stress supérieur au stress usuel.

Plusieurs plans d'expérience sont utilisés en ALT. Nous pouvons en citer :

- **Premier plan d'expérience**

Notons par $z^{(0)} = (z_{00}, z_{01}, \dots, z_{0m})$, $z_{00} = 1$, le stress usuel. Généralement, nous prenons $m = 1$ (stress unidimensionnel), ou $m = 2$ (stress bidimensionnel).

Soient $z^{(1)}, \dots, z^{(k)}$ des stress constants dans le temps tels que :

$$z^{(0)} < z^{(1)} < \dots < z^{(k)}$$

avec $z_i = (z_{i0}, z_{i1}, \dots, z_{im})^T \in E_m$, $z_{i0} = 1$. Le stress usuel $z^{(0)}$ n'est pas utilisé durant l'expérience.

Selon le premier plan d'expérience, k groupes d'unités sont testés. Le $i^{\text{ème}}$ groupe de n_i unités, $\sum_{i=1}^k n_i = n$, est testé sous le stress $z^{(i)}$. Les données peuvent être complètes ou censurées à droite.

- **Deuxième plan d'expérience**

En deuxième plan d'expérience, les step-stress sont utilisés (voir la section 2.1). Si la fonction $r(z)$ est paramétrisée comme suit :

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

Alors la fonction de survie de l'équation (2.4) sous le stress $z(\cdot)$ de la forme (2.5) peut s'écrire :

$$S_{z(\cdot)}(t) = S_0 \left\{ 1_{\{i>1\}} \sum_{j=1}^{i-1} e^{-\beta^T z_j} (t_j - t_{j-1}) + e^{-\beta^T z_i} (t - t_{i-1}) \right\}, t \in [t_{i-1}, t_i[, i = 1, \dots, m.$$

Troisième plan d'expérience

Dans ce plan, deux groupes d'unités sont testés. Le premier groupe d'unités n_1 est testé sous un stress accéléré constant z_1 et le second groupe d'unités n_2 est testé sous un step-stress z_2 ; le stress z_1 jusqu'au moment t_1 , ensuite sous le stress z_0 jusqu'à l'instant t_2 c'est à dire :

$$z_2(\omega) = \begin{cases} z_1, & 0 \leq \omega \leq t_1 \\ z_0, & t_1 \leq \omega \leq t_2 \end{cases}$$

Les unités utilisent beaucoup de leurs ressources jusqu'au moment t_1 sous le stress accéléré z_1 , donc après les défaillances qui se produisent dans l'intervalle $[t_1, t_2]$ sont pareilles au stress habituel ou normal.

La fonction de survie dans le modèle AFT dans ce cas, est

$$S_{z_1}(u) = S_{z_0}(ru),$$

où $r = r(z_1)/r(z_0)$, et

$$S_{z_2(\cdot)}(u) = \begin{cases} S_{z_0}(ru), & 0 \leq u \leq t_1 \\ S_{z_0}(ru)(rt_1 + u - t_1), & t_1 \leq u \leq t_2 \end{cases}$$

où

$$S_{z_2(\cdot)}(t) = S_{z_0}(r(u \wedge t_1) + (u - t_1) \vee 0)$$

et

$$a \wedge b = \min(a, b) \quad \text{et} \quad a \vee b = \max(a, b).$$

Quatrième plan d'expérience

Ce plan d'expérience est utilisé lorsque la distribution du taux de défaillance est exponentielle. Pour ce plan k groupes d'unités sont observés. Le $i^{\text{ème}}$ groupe d'unités n_i est testé sous un stress constant unidimensionnel $z^{(i)}$ jusqu'à la $r_i^{\text{ème}}$ défaillance ($r_i \leq n_i$) (censure de type II). Les moments de défaillance du $i^{\text{ème}}$ groupe sont $T_{i1} \leq \dots \leq T_{ir_i}$, où $i = 1, \dots, k$.

Remarque : Nous utilisons seulement les deux premiers plans d'expérience.

2.2 Construction du modèle AFT-IWG

Pour un modèle AFT dont la distribution de base est une inverse Weibull généralisée, la fonction de survie de base $S_0(t)$ est donnée par :

$$S_0(t) = 1 - \exp\left[-\gamma\left(\frac{\alpha}{t}\right)^\nu\right], \quad t > 0, \alpha > 0, \nu > 0, \gamma > 0$$

La fonction de distribution cumulative de AFT-IWG est déduite comme suit :

$$\begin{aligned} F(t) &= 1 - S_0(\exp\{-\beta^T z\} t) \\ &= \exp[-\gamma\alpha^\nu \exp\{\nu\beta^T z\} t^{-\nu}], \end{aligned}$$

où α, ν et γ sont les paramètres de la distribution IWG,

$\beta = (\beta_0, \dots, \beta_m)^T$ sont les paramètres de la régression du modèle AFT,
et

$z = (1, z_1, \dots, z_m)^T$ sont les covariables représentant les stress éventuels.

Ainsi, la fonction de survie pour le modèle AFT-IWG est obtenue comme suit :

$$\begin{aligned} S(t) &= S_0(\exp \{-\beta^T z\} t) \\ &= 1 - \exp[-\gamma \alpha^\nu \exp \{\nu \beta^T z\} t^{-\nu}] \end{aligned}$$

Les fonctions risque et risque cumulé sont alors, respectivement :

$$\begin{aligned} \lambda(t) &= -S'(t)/S(t) \\ &= \gamma \nu \alpha^\nu \exp[\nu \beta^T z] t^{-(\nu+1)} \exp[-\gamma \alpha^\nu \exp \{\nu \beta^T z\} t^{-\nu}] \\ &\quad [1 - \exp[-\gamma \alpha^\nu \exp \{\nu \beta^T z\} t^{-\nu}]]^{-1} \end{aligned}$$

et

$$\begin{aligned} \Lambda(t) &= -\ln S(t) \\ &= -\ln(1 - \exp[-\gamma \alpha^\nu \exp \{\nu \beta^T z\} t^{-\nu}]) \end{aligned}$$

Nous présentons dans la figure Fig. 12, les différentes formes du taux de hasard (croissante, décroissante, en cloche et en baignoire) de l'AFT-IWG en variant la valeur du paramètre de forme ν .

Comme nous le voyons sur les graphes, la distribution AFT-IWG peut avoir différentes formes de la fonction risque :

-Si $\nu < 1$, la fonction risque diminue continuellement, ce qui représente les pannes précoce.

-Si $\nu > 1$, la fonction risque augmente continuellement, ce qui représente le vieillissement.

-Si $\nu = 1$ et $\gamma = 1$, la fonction risque représente les pannes accidentelles.

2.3 Estimation du maximum de vraisemblance en cas de données censurées

Soit T_i est une variable aléatoire selon une distribution de la loi AFT-IWG, avec le vecteur de paramètres $\theta = (\alpha, \nu, \gamma, \beta_0, \beta_1, \dots, \beta_m)^T$.

Supposons que les données se composent de n observations indépendantes $t_i = \min(T_i, C_i)$ pour $i = 1, \dots, n$. La distribution de C_i ne dépend pas des paramètres inconnus de T_i .

La fonction de vraisemblance dans le cas de la censure

$$L(t, \theta) = \prod_{i=1}^n f(t_i, \theta)$$

est équivalente à

$$L(t, \theta) = \prod_{i=1}^n \lambda^{\delta_i}(t_i, \theta) S(t_i, \theta), \quad \delta_i = 1_{\{T_i \leq C_i\}}$$

Dans le cas de la distribution AFT-IWG, nous avons :

$$\begin{aligned} L(t, \theta) &= \prod_{i=1}^n [\nu \gamma \alpha^\nu t_i^{-(\nu+1)} \exp[\nu \beta^T z_i] \exp[-\gamma \alpha^\nu \exp[\nu \beta^T z_i] t_i^{-\nu}] \\ &\quad [1 - \exp[-\gamma \alpha^\nu \exp[\nu \beta^T z_i] t_i^{-\nu}]]^{-1}]^{\delta_i} [1 - \exp[-\gamma \alpha^\nu \exp[\nu \beta^T z_i] t_i^{-\nu}]] \end{aligned}$$

La fonction de la log-vraisemblance dans ce cas est :

$$\begin{aligned}
\ell(t, \theta) &= \sum_{i=1}^n [\delta_i [\ln(\nu) + \ln(\gamma) + \nu \ln(\alpha) - (\nu + 1) \ln(t_i) + \nu \beta^T z_i - \gamma \alpha^\nu \exp [\nu \beta^T z_i] t_i^{-\nu} \\
&\quad - \ln (1 - \exp [-\gamma \alpha^\nu \exp [\nu \beta^T z_i] t_i^{-\nu}])] + \ln (1 - \exp [-\gamma \alpha^\nu \exp [\nu \beta^T z_i] t_i^{-\nu}])] \\
&= r (\log(\nu) + \log(\gamma) + \nu \log(\alpha) + \nu \beta^T z_i) - (\nu + 1) \sum_{i \in F} \log(t_i) \\
&\quad - \gamma \alpha^\nu \sum_{i \in F} \exp \{\nu \beta^T z_i\} t_i^{-\nu} + \sum_{i \in C} \log (1 - \exp [-\gamma \alpha^\nu \exp \{\nu \beta^T z_i\} t_i^{-\nu}])
\end{aligned}$$

où r est le nombre de défaillances, F et C désignent respectivement les ensembles d'observations non censurées et censurées.

Après de longs calculs, nous obtenons les fonctions de score pour les paramètres $\alpha, \nu, \gamma, \beta_0, \beta_1, \dots, \beta_m$:

$$\begin{aligned}
\frac{\partial \ell(t; \theta)}{\partial \alpha} &= \frac{r\nu}{\alpha} - \nu \gamma \alpha^{\nu-1} \sum_{i \in F} \exp [\nu \beta^T z_i] t_i^{-\nu} \\
&\quad + \nu \gamma \alpha^{\nu-1} \sum_{i \in C} [\exp [\beta^T z_i] t_i^{-\nu} \exp [-\gamma \alpha^\nu \exp [\nu \beta^T z_i] t_i^{-\nu}]] \\
&\quad \times [1 - \exp [-\gamma \alpha^\nu \exp [\nu \beta^T z_i] t_i^{-\nu}]]^{-1} \\
\frac{\partial \ell(t; \theta)}{\partial \nu} &= \frac{r}{\nu} + r \ln(\alpha) - \sum_{i \in F} (\log(t_i) + \beta^T z_i) \\
&\quad - \gamma \sum_{i \in F} \ln (\alpha \exp [\beta^T z_i] t_i^{-1}) \alpha^\nu \exp [\nu \beta^T z_i] t_i^{-\nu} \\
&\quad + \gamma \alpha^\nu \sum_{i \in C} \ln (\alpha \exp [\beta^T z_i] t_i^{-1}) [\exp [\beta^T z_i] t_i^{-\nu} \exp [-\gamma \alpha^\nu \exp [\nu \beta^T z_i] t_i^{-\nu}]] \\
&\quad \times [1 - \exp [-\gamma \alpha^\nu \exp [\nu \beta^T z_i] t_i^{-\nu}]]^{-1}
\end{aligned}$$

$$\begin{aligned}
\frac{\partial \ell(t; \theta)}{\partial \gamma} &= \frac{r}{\gamma} - \alpha^\nu \sum_{i \in F} \exp [\nu \beta^T z_i] t_i^{-\nu} \\
&\quad + \alpha^\nu \sum_{i \in C} [\exp [\beta^T z_i] t_i^{-\nu} \exp [-\gamma \alpha^\nu \exp [\nu \beta^T z_i] t_i^{-\nu}]] \\
&\quad \times [1 - \exp [-\gamma \alpha^\nu \exp [\nu \beta^T z_i] t_i^{-\nu}]]^{-1} \\
\\
\frac{\partial \ell(t; \theta)}{\partial \beta_0} &= r\nu - \gamma \alpha^\nu \nu \sum_{i \in F} \exp [\nu \beta^T z_i] t_i^{-\nu} \\
&\quad + \gamma \alpha^\nu \nu \sum_{i \in C} [\exp [\beta^T z_i] t_i^{-\nu} \exp [-\gamma \alpha^\nu \exp [\nu \beta^T z_i] t_i^{-\nu}]] \\
&\quad \times [1 - \exp [-\gamma \alpha^\nu \exp [\nu \beta^T z_i] t_i^{-\nu}]]^{-1} \\
\\
\frac{\partial \ell(t; \theta)}{\partial \beta_1} &= \sum_{i \in F} \nu z_{i1} - \gamma \alpha^\nu \nu z_{i1} \exp [\nu \beta^T z_i] t_i^{-\nu} \\
&\quad + \gamma \alpha^\nu \nu \sum_{i \in C} z_i [\exp [\beta^T z_i] t_i^{-\nu} \exp [-\gamma \alpha^\nu \exp [\nu \beta^T z_i] t_i^{-\nu}]] \\
&\quad \times [1 - \exp [-\gamma \alpha^\nu \exp [\nu \beta^T z_i] t_i^{-\nu}]]^{-1} \\
\\
&\dots \\
\frac{\partial \ell(t; \theta)}{\partial \beta_m} &= \sum_{i \in F} \nu z_{im} - \gamma \alpha^\nu \nu z_{im} \exp [\nu \beta^T z_i] t_i^{-\nu} \\
&\quad + \gamma \alpha^\nu \nu \sum_{i \in C} z_i [\exp [\beta^T z_i] t_i^{-\nu} \exp [-\gamma \alpha^\nu \exp [\nu \beta^T z_i] t_i^{-\nu}]] \\
&\quad \times [1 - \exp [-\gamma \alpha^\nu \exp [\nu \beta^T z_i] t_i^{-\nu}]]^{-1}
\end{aligned}$$

2.4 Estimation et intervalle de confiance de la fonction de survie

2.4.1 Estimation de la fonction de survie

Si $\hat{\alpha}, \hat{\nu}, \hat{\gamma}$ et $\hat{\beta}$ sont les estimateurs de maximum de vraisemblance des paramètres du modèle AFT-GIW, alors la fonction de survie sous un stress d'accélération constant $z \in E_0$ est :

$$\hat{S}_z(t) = 1 - \exp \left[-\hat{\gamma} \hat{\alpha}^{\hat{\nu}} \exp \left\{ \hat{\nu} \hat{\beta}^T z \right\} t^{-\hat{\nu}} \right]$$

Et la fonction de survie estimée sous un stress normal ou usuel quand $z = z^{(0)}$ peut être écrite comme suit :

$$\hat{S}_{z^{(0)}}(t) = 1 - \exp \left[-\hat{\gamma} \hat{\alpha}^{\hat{\nu}} \exp \left\{ \hat{\nu} \hat{\beta}^T z^{(0)} \right\} t^{-\hat{\nu}} \right]$$

2.4.2 Intervalle de confiance de la fonction de survie

En utilisant les propriétés des estimateurs du maximum du vraisemblance, sous le stress usuel, pour un risque α , l'intervalle de confiance pour la fonction de survie $\hat{S}_{z^{(0)}}(t)$ est :

$$\left(1 + \frac{1 - \hat{S}_{z^{(0)}}(t)}{\hat{S}_{z^{(0)}}(t)} \exp \left[\pm \hat{\sigma} Q_{z^{(0)}} \omega_{1-\frac{\alpha}{2}} \right] \right)^{-1}$$

Où ω_α est le α -quantile de la distribution normale et

$$\hat{\sigma}^2 Q_{z^{(0)}} = \frac{J_g^T(\hat{\theta}) I_g^{-1}(\hat{\theta}) J_g(\hat{\theta})}{\left(\hat{S}_{z^{(0)}}(t) \right)^2 \left(1 - \hat{S}_{z^{(0)}}(t) \right)^2}$$

$$\begin{aligned}
J_g^T(\hat{\theta}) I_g^{-1}(\hat{\theta}) J_g(\hat{\theta}) &= \left(\frac{\partial S}{\partial \beta_0}, \dots, \frac{\partial S}{\partial \beta_m}, \frac{\partial S}{\partial \alpha}, \frac{\partial S}{\partial \nu}, \frac{\partial S}{\partial \gamma} \right) \begin{pmatrix} I_{00} & \cdot & \cdot & \cdot & I_{0(m+3)} \\ \vdots & & & & \vdots \\ I_{i0} & \cdot & \cdot & \cdot & I_{i(m+3)} \\ \vdots & & & & \vdots \\ I_{(m+3)0} & \cdot & \cdot & \cdot & I_{(m+3)(m+3)} \end{pmatrix} \begin{pmatrix} \frac{\partial S}{\partial \beta_0} \\ \vdots \\ \frac{\partial S}{\partial \beta_m} \\ \frac{\partial S}{\partial \alpha} \\ \frac{\partial S}{\partial \nu} \\ \frac{\partial S}{\partial \gamma} \end{pmatrix} \\
&= \left(\frac{\partial S}{\partial \beta_0}, \dots, \frac{\partial S}{\partial \beta_m}, \frac{\partial S}{\partial \alpha}, \frac{\partial S}{\partial \nu}, \frac{\partial S}{\partial \gamma} \right) \begin{pmatrix} \sum_{j=0}^{m+3} I_{0j} \frac{\partial S}{\partial \beta_j} \\ \sum_{j=0}^{m+3} I_{1j} \frac{\partial S}{\partial \beta_j} \\ \vdots \\ \sum_{j=0}^{m+3} I_{(m+3)j} \frac{\partial S}{\partial \beta_j} \end{pmatrix} \\
&= \sum_{k=0}^{m+3} \sum_{j=0}^{m+3} \frac{\partial S}{\partial \beta_k} I_{kj} \frac{\partial S}{\partial \beta_j}
\end{aligned}$$

Où il peut être écrit sous la forme simple :

$$\widehat{\sigma}^2 Q_{z^{(0)}} = \frac{1}{\left(\widehat{S}_{z^{(0)}}(t)\right)^2 \left(1 - \widehat{S}_{z^{(0)}}(t)\right)^2} \sum_{k=0}^{m+3} \sum_{j=0}^{m+3} a_k(t; \widehat{\alpha}, \widehat{\nu}, \widehat{\gamma}, \widehat{\beta}) I_{kj}(\widehat{\alpha}, \widehat{\nu}, \widehat{\gamma}, \widehat{\beta}) a_j(t; \widehat{\alpha}, \widehat{\nu}, \widehat{\gamma}, \widehat{\beta})$$

Où $a_k(t; \widehat{\alpha}, \widehat{\nu}, \widehat{\gamma}, \widehat{\beta})$ et $a_j(t; \widehat{\alpha}, \widehat{\nu}, \widehat{\gamma}, \widehat{\beta})$ sont les dérivées partielles de la fonction de survie par rapport aux paramètres du modèle AFT-IWG tels que :

$$\begin{aligned}
\frac{\partial S}{\partial \alpha} &= \gamma \nu \alpha^{\nu-1} \exp[\nu \beta^T z] t^{-\nu} \exp[-\gamma \alpha^\nu \exp(\nu \beta^T z) t^{-\nu}] \\
\frac{\partial S(t, \theta)}{\partial \nu} &= \gamma \alpha^{\nu-1} t^{\nu-1} \exp[\nu \beta^T z] [\nu + \alpha t^{-1} \beta^T z] \exp[-\gamma \alpha^\nu \exp(\nu \beta^T z) t^{-\nu}]
\end{aligned}$$

$$\begin{aligned}
\frac{\partial S(t, \theta)}{\partial \gamma} &= \alpha^\nu \exp [\nu \beta^T z] t^{-\nu} \exp [-\gamma \alpha^\nu \exp (\nu \beta^T z) t^{-\nu}] \\
\frac{\partial S(t, \theta)}{\partial \beta_0} &= \gamma \nu \alpha^{\nu-1} t^{-\nu} \exp [\nu \beta^T z] \exp [-\gamma \alpha^\nu \exp (\nu \beta^T z) t^{-\nu}] \\
\frac{\partial S(t, \theta)}{\partial \beta_1} &= \gamma \nu z_1 \alpha^{\nu-1} t^{-\nu} \exp [\nu \beta^T z] \exp [-\gamma \alpha^\nu \exp (\nu \beta^T z) t^{-\nu}] \\
&\vdots \\
\frac{\partial S(t, \theta)}{\partial \beta_m} &= \gamma \nu z_m \alpha^{\nu-1} t^{-\nu} \exp [\nu \beta^T z] \exp [-\gamma \alpha^\nu \exp (\nu \beta^T z) t^{-\nu}]
\end{aligned}$$

Et $I_{kj}(\widehat{\alpha}, \widehat{\nu}, \widehat{\gamma}, \widehat{\beta})$ sont les éléments de la matrice $I^{-1}(\widehat{\alpha}, \widehat{\nu}, \widehat{\gamma}, \widehat{\beta})$. La matrice de Fisher $I_{kj}(\alpha, \nu, \gamma, \beta)$ est estimée par :

$$I(\widehat{\alpha}, \widehat{\nu}, \widehat{\gamma}, \widehat{\beta}) = -\frac{\partial^2 l(\widehat{\alpha}, \widehat{\nu}, \widehat{\gamma}, \widehat{\beta})}{\partial \theta_i \partial \theta_j}, \quad \theta = (\alpha, \nu, \gamma, \beta_0, \beta_1, \dots, \beta_m)^T$$

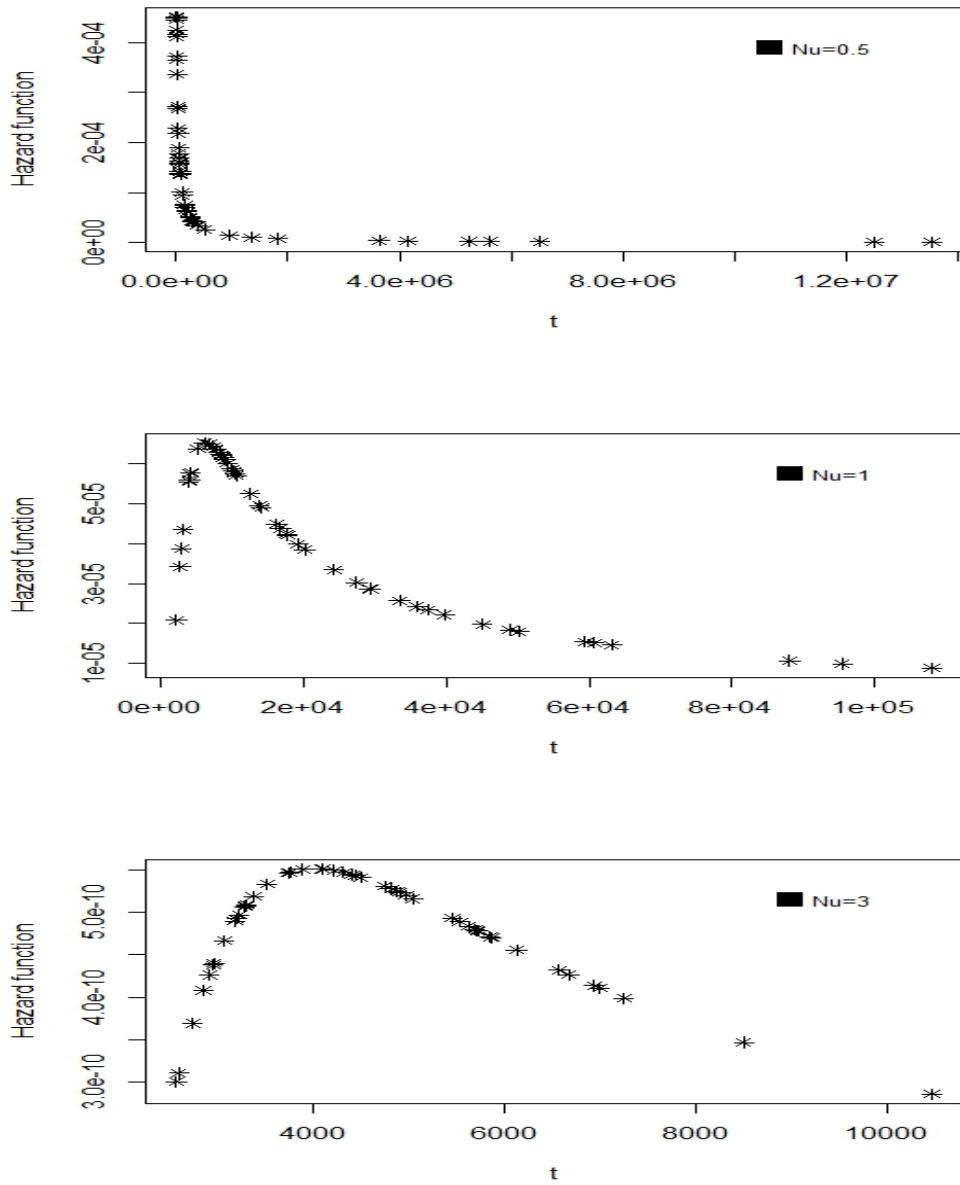


Fig. 12. Taux de hasard de l'AFT-IWG pour $\alpha = 2$, $\gamma = 5$, $\beta_0 = 6$,
 $\beta_1 = 0.82$ avec $\nu = (0.5, 1, 3)$.

Chapitre 3

Tests d'ajustement pour données censurées

Un test d'ajustement du Chi-deux peut être la première étape d'une analyse des données plus approfondie. C'est une technique utile pour voir si les données observées sont représentatives d'un modèle ou d'une distribution. Le fameux test d'ajustement du Chi-deux a été utilisé dans presque tous les domaines de la recherche depuis sa découverte par Karl Pearson (1900). La distribution limite de la statistique de Pearson est un chi-deux si la distribution est discrète et si les paramètres sont connus. Mais, si les paramètres sont estimés par la méthode du minimum du Chi-deux ou la méthode d'estimation du maximum de vraisemblance pour les données groupées, la distribution limite ne sera plus un chi-deux, et elle dépendra même de la méthode d'estimation utilisée (Cramer, 1946).

Pour surmonter le problème de l'estimation de paramètres pour les données non groupées et la distribution limite de la statistique, de nombreuses modifications ont été apportées sur le classique test de Pearson et une variété de procédures ont été mises en oeuvre dans de nombreux logiciels.

Nikulin (1973a, 1973b, 1973c), Rao et Robson (1974) ont proposé indépendamment une modification de la statistique de Pearson dans le cas où les données sont complètes. La statistique proposée connue sous le nom "la statistique de Nikulin-Rao-Robson (NRR)" (Drost, 1988 ; Van der Vaart, 1998 ; Voinov et al, 2008), et dont la distribution limite suit un chi-deux, est très utilisée.

3.1 Tests d'ajustement pour des données censurées à droite

Toutefois, en analyse de survie et en fiabilité, nous rencontrons souvent des observations incomplètes (censurées), dans cette situation les méthodes habituelles ne sont plus valables. Plusieurs tests d'ajustement ont été proposés pour des données censurées aléatoires à droite.

Chen (1975) avait construit un test du type Chi-deux, Habib et Thomas (1986) ont développé une statistique du type Chi-deux où l'estimateur de Kaplan-Meier $\hat{S}_n(x)$ est comparé à l'estimateur paramétrique $S(x, \hat{\theta}_n)$, où $\hat{\theta}_n$ est l'estimateur du maximum de vraisemblance de θ_n . Akritas (1988) a proposé une statistique du type Chi-deux basée sur l'idée de comparer les nombres de pannes observées et prévues dans chaque classe. Hjort (1990) a élaboré une statistique du type Chi-deux pour tester la validité du modèle paramétrique pour les données de vie basée sur le processus du hasard cumulatif. Kim (1993) a également proposé un test du type Chi-deux basé sur l'estimateur non paramétrique de type produit-limite. Nikulin et Solev (1999) ont construit un test de type Chi-deux pour les données censurées doubles.

Dans cette section, nous présentons d'abord le test de Habib et Thomas (1986), ensuite, nous construisons pour le modèle AFT-IWG, introduit dans le chapitre précédent, un test d'ajustement du type du Chi-deux modifié basé sur Y_n^2 , une modification de la statistique NRR proposée par Bagdonavčius et Nikulin (2011) dans le cas de données censurées, basée sur les estimateurs du maximum de vraisemblance et les données non regroupées.

3.1.1 Test de Habib et Thomas

Ce test est bien adapté pour les données censurées aléatoires à droite où l'estimateur de Kaplan-Meier $\hat{S}_n(t)$ est comparé à l'estimateur paramétrique $S(t, \hat{\theta}_n)$, ($\hat{\theta}_n$ est l'estimateur du maximum de vraisemblance de θ_n).

Habib et Thomas en 1986 ont démontré que :

$$\sqrt{n} \left(\hat{S}_n(t) - S(t, \hat{\theta}_n) \right)$$

converge vers un processus Gaussien, sous l'hypothèse H_0 .

On partage la droite réelle en r intervalles : I_1, I_2, \dots, I_r mutuellement disjoints par les points :

$$0 = t_0 < t_1 < \dots < t_{r-1} < t_r = +\infty.$$

Considérons le vecteur :

$$\hat{Z}_n = \sqrt{n} \left(\hat{S}_n - S_{\hat{\theta}_n} \right),$$

où

$$\hat{S}_n = (\hat{S}_n(t_1), \hat{S}_n(t_2), \dots, \hat{S}_n(t_{r-1}))^T \quad \text{et} \quad S_{\hat{\theta}_n} = (S(t_1, \hat{\theta}_1), S(t_2, \hat{\theta}_2), \dots, S(t_{r-1}, \hat{\theta}_n))^T.$$

Sous les suppositions suivantes :

- 1) $f(t, \theta)$ et $F(t, \theta)$ sont deux fois continûment dérivables, par rapport à θ ,
- 2) la matrice d'information de Fisher $I(\theta)$ est définie positive, et continue par rapport à θ ,

$$I_{ij} = - \int \frac{\partial^2 \ln f(t, \theta)}{\partial \theta_i \partial \theta_j} H(t) f(t, \theta) dt - \int \frac{\partial^2 \ln S(t, \theta)}{\partial \theta_i \partial \theta_j} h(t) S(t, \theta) dt, \quad i, j = 1, 2, 3.$$

3) l'estimateur du Maximum de vraisemblance $\hat{\theta}_n$ de θ existe et $\sqrt{n}-$ convergent avec

$$\sqrt{n}(\hat{\theta}_n - \theta) = I^{-1}W_n + o_p(1),$$

où

$$W_n = n^{-1/2} \sum_{i=1}^n \frac{\partial \ln g(t_i, \delta_i, \theta)}{\partial \theta},$$

et g est la densité jointe de la distribution de (t, δ) .

Soit

$$B = B(\theta) = \left[\frac{\partial F(t_i, \theta)}{\partial \theta_j} \right]_{(r-1) \times 2},$$

et

$$V = V(\theta) = [Cov(Z(t_i), Z(t_j))]_{(r-1) \times (r-1)},$$

où

$$Cov(Z(t_i), Z(t_j)) = S(t_i, \theta)S(t_j, \theta) \int_0^{t_i \wedge t_j} \frac{f(t, \theta)}{H(t)F^2(t, \theta)} dt.$$

Pour tester H_0 , selon laquelle la variable aléatoire T suit une loi de probabilité F_0 , Habib et Thomas ont construit le test du Chi-deux modifié de Pearson qui a la forme quadratique suivante :

$$\hat{Y}_n^2(\hat{\theta}_n) = \hat{Z}_n^T \hat{\Sigma}^-(\hat{\theta}_n) \hat{Z}_n,$$

où la matrice $\hat{\Sigma}$ est l'estimateur de la matrice de covariance Σ et Σ^- est son inverse généralisée, de sorte que :

$$\Sigma(\theta) = V(\theta) - B(\theta)I^{-1}(\theta)B^T(\theta), \quad \text{rang}\Sigma = r - 1.$$

Le comportement asymptotique de la statistique $\hat{Y}_n^2(\hat{\theta}_n)$ est donné par le théorème suivant :

Théorème [Habib et Thomas (1986)]
 Sous les conditions de régularité ci-dessus, et pour n suffisamment grand :

$$\lim_{n \rightarrow \infty} \mathbf{P}(\hat{Y}_n^2(\hat{\theta}_n)) \leq x | H_0 = \mathbf{P}(\chi_{r-1}^2 \leq x).$$

Notons qu'en l'absence de la censure, la statistique \hat{Y}_n^2 se réduit à celle de la statistique NRR. Le cas de données avec censure double a aussi été considéré (voir Nikulin et Solev (1999)).

Test de Habib et Thomas pour la distribution IWG

Comme nous l'avons précédemment mentionné, la statistique du test de Habib et Thomas est :

$$\hat{Y}_n^2(\hat{\theta}_n) = \widehat{\hat{Z}_n^T \sum (\hat{\theta}_n) \hat{Z}_n}$$

Les éléments de la statistique $\hat{Y}_n^2(\hat{\theta}_n)$ de ce test dans le cas de la distribution IWG sont :

$$\sum(\theta) = V(\theta) - B(\theta)I^{-1}(\theta)B(\theta)$$

où

$$B(\theta) = \left[\frac{\partial F(t_i, \theta)}{\partial \theta_j} \right] \quad \theta = (\alpha, \beta, \gamma) \quad \text{et} \quad j = 1, 2, 3.$$

La fonction de densité cumulative de l'IWG est :

$$F(t_i, \theta) = \exp \left[-\gamma \left(\frac{\alpha}{t_i} \right)^\beta \right]$$

Ses dérivées sont données par :

$$\begin{aligned} \frac{\partial F(t_i, \theta)}{\partial \alpha} &= -\beta \gamma \alpha^{\beta-1} t_i^{-\beta} \exp \left[-\gamma \left(\frac{\alpha}{t_i} \right)^\beta \right] \\ \frac{\partial F(t_i, \theta)}{\partial \beta} &= -\gamma \alpha^\beta t_i^{-\beta} \ln \alpha \exp \left[-\gamma \left(\frac{\alpha}{t_i} \right)^\beta \right] \\ \frac{\partial F(t_i, \theta)}{\partial \gamma} &= -\alpha^\beta t_i^{-\beta} \exp \left[-\gamma \left(\frac{\alpha}{t_i} \right)^\beta \right] \end{aligned}$$

La matrice $V_{(r-1) \times (r-1)}$ est définie par la formule suivante :

$$V = V(\theta) = [Cov(Z(t_i), Z(t_j))]$$

Où

$$\begin{aligned} Z_n(t_i) &= \sqrt{n} \left(\widehat{S}_n(t_i) - \widehat{S}_n(t_i, \widehat{\theta}_n) \right) \\ &= \sqrt{n} \left(\left[1 - \exp \left[-\gamma \left(\frac{\alpha}{t_i} \right)^\beta \right] \right] - \left[1 - \exp \left[-\widehat{\gamma} \left(\frac{\widehat{\alpha}}{t_i} \right)^{\widehat{\beta}} \right] \right] \right) \end{aligned}$$

3.1.2 Test de Bagdonavičius et Nikulin

Dans cette section, nous présentons le test proposé par Bagdonavičius et Nikulin, pour tester des hypothèses paramétriques lorsque les données sont censurées à droite.

Soit l'échantillon

$$(X_1, \delta_1), (X_2, \delta_2), \dots, (X_n, \delta_n) \quad (3.12)$$

où

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

Supposons que les taux de panne T_1, T_2, \dots, T_n sont n variables aléatoires indépendantes et identiquement distribuées. La fonction de densité de T_i appartient à une famille paramétrique $\{f(t, \theta), \theta \in \Theta \subset \mathbf{R}^m\}$.

Notons par :

$$\Lambda(t, \theta) = -\ln S(t, \theta) = \int_0^t \lambda(u, \theta) du$$

la fonction de hasard cumulative.

La fonction de la log-vraisemblance est :

$$\ell(\theta) = \sum_{i=1}^n \delta_i \ln \lambda(X_i, \theta) + \sum_{i=1}^n \ln S(X_i, \theta) \quad (3.13)$$

l'estimateur du maximum de vraisemblance $\hat{\theta}$ de θ vérifie :

$$\dot{\ell}(\hat{\theta}) = \mathbf{0}$$

où

$$\dot{\ell}(\theta) = \frac{\partial}{\partial \theta} \ell(\theta) = \left(\frac{\partial}{\partial \theta_1} \ell(\theta), \dots, \frac{\partial}{\partial \theta_m} \ell(\theta) \right)^T \quad (3.14)$$

La matrice d'information de Fisher est :

$$I(\theta) = -E_{\theta} \ddot{\ell}(\theta) \quad (3.15)$$

où

$$\ddot{\ell}(\theta) = \sum_{i=1}^n \delta_i \frac{\partial^2}{\partial \theta^2} \ln \lambda(X_i, \theta) - \sum_{i=1}^n \frac{\partial^2}{\partial \theta^2} \Lambda(X_i, \theta) \quad (3.16)$$

Si θ_0 est la vraie valeur de θ et sous les conditions de régularité, on a :

$$\hat{\theta} \xrightarrow{P} \theta_0; \quad \sqrt{n}(\hat{\theta} - \theta_0) \xrightarrow{P} N(\mathbf{0}_m, i^{-1}(\theta_0)), \quad 0_m = (0, 0, \dots, 0)^T;$$

$$\frac{1}{\sqrt{n}} \dot{\ell}(\theta_0) \xrightarrow{d} N(\mathbf{0}_m, i(\theta_0)); \quad -\frac{1}{n} \ddot{\ell}(\hat{\theta}) \xrightarrow{P} i(\theta_0),$$

L'échantillon (3.12) peut être représenté en termes de processus de comptage comme suit :

$$(N_1(t), Y_1(t), t \geq 0), \dots, (N_n(t), Y_n(t), t \geq 0) \quad (3.17)$$

où

$$N_i(t) = \mathbf{1}_{\{X_i \leq t, \delta_i=1\}}, \quad Y_i(t) = \mathbf{1}_{\{0 < t \leq X_i\}}.$$

Posons

$$N(t) = \sum_{i=1}^n N_i(t) \quad \text{et} \quad Y(t) = \sum_{i=1}^n Y_i(t) \quad (3.18)$$

Écrivons les expressions des équations (3.13), (3.14), (3.15) et (3.16) en termes des processus stochastiques N_i et Y_i . Les trajectoires du processus de comptage $N_i(t)$ ont la forme :

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

Si $\delta_i = 1$, et $N_i(t) = 0$, si $\delta_i = 0$, les trajectoires du processus de comptage Y_i ont la forme :

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

En utilisant ces processus, on obtient les relations suivantes :

$$\int_0^\infty \ln \lambda(u, \theta) dN_i(u) = \begin{cases} \ln \lambda(X_i, \theta), & \delta_i = 1, \\ 0, & \delta_i = 0. \end{cases} = \delta_i \ln \lambda(X_i, \theta) \quad (3.19)$$

et

$$\int_0^\infty Y_i(u) \lambda(u) du = \int_0^{X_i} \lambda(u) du = -\ln S(X_i, \theta) \quad (3.20)$$

Alors les équations (3.13), (3.14), (3.15) et (3.16) s'écrivent :

$$\ell(\theta) = \int_0^\tau \{\ln \lambda(u, \theta) dN(u) - Y(u)\lambda(u, \theta)\} du \quad (3.21)$$

$$\dot{\ell}(\theta) = \int_0^\tau \frac{\partial}{\partial \theta} \ln \lambda(u, \theta) \{dN(u) - Y(u)\lambda(u, \theta)du\} \quad (3.22)$$

$$\ddot{\ell}(\theta) = \int_0^\tau \frac{\partial^2}{\partial \theta^2} \ln \lambda(u, \theta) dN(u) - \int_0^\tau \frac{\partial^2}{\partial \theta^2} \lambda(u, \theta) Y(u) du \quad (3.23)$$

La matrice d'information de Fisher est :

$$I(\theta) = -E_\theta \ddot{\ell}(\theta) = \mathbf{E}_\theta \sum_{i=1}^n \int_0^\infty \frac{\partial}{\partial \theta} \ln \lambda(u, \theta) \left(\frac{\partial}{\partial \theta} \ln \lambda(u, \theta) \right)^T \lambda(u, \theta) Y_i(u) du \quad (3.24)$$

Sous les conditions suivantes [Hjort (1990)] :

1. Il existe un voisinage Θ_0 de θ_0 la vraie valeur de θ tel que pour tout n et $\theta \in \Theta_0$, et pour tout $t \in [0, \tau]$, les dérivées partielles d'ordre 1, 2 et 3 de $\lambda(t, \theta)$ par rapport à θ existent et sont continues et bornées. On suppose aussi que la fonction de log-vraisemblance est trois fois différentiable par rapport à $\theta \in \Theta_0$.
2. $\lambda(t, \theta)$ est bornée dans $[0, \tau] \times \Theta_0$.
3. Il existe une fonction positive $y(t)$ telle que :

$$\sup_{t \in [0, \tau]} |Y(t)/n - y(t)| \xrightarrow{P} 0$$

4. La matrice $i(\theta_0) = \lim_{n \rightarrow \infty} \frac{I(\theta_0)}{n}$ (la limite existe sous les conditions 1-3) est définie positive.

Supposons que les processus N_i et Y_i sont observés à un temps fini τ , après on partage l'intervalle $[0, \tau]$ en $k > s$ sous-intervalles

$$I_j = (a_{j-1}, a_j], \quad a_0 = 0, \quad a_k = \tau,$$

et notons par :

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

le nombre des défaillances observées dans le j^{eme} intervalle, $j = 1, 2, \dots, k$.

On a besoin d'estimer le nombre de défaillances théorique dans l'intervalle I_j sous l'hypothèse H_0 ; en prenant en compte l'équation :

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

$$E_j = \int_{a_{j-1}}^{a_j} \lambda(u, \hat{\theta}) Y(u) du = \sum_{i:X_i > a_{j-1}} (\Lambda(a_j \wedge X_i, \hat{\theta}) - \Lambda(a_{j-1}, \hat{\theta}))$$

$$\text{où } a \wedge b = \min(a, b).$$

On peut prévoir d'observer le nombre de défaillance

$$e_j = \int_{a_{j-1}}^{a_j} \lambda(u, \hat{\theta}) Y(u) du$$

Ici $\hat{\theta}$ est l'estimateur du maximum de vraisemblance du paramètre θ .

Le test d'ajustement est basé sur la statistique :

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

Le vecteur Z est donné par :

$$Z = (Z_1, Z_2, \dots, Z_k)^T, \quad Z_j = \frac{1}{\sqrt{n}}(U_j - E_j)$$

l'inverse généralisée de la matrice \hat{V} , basé sur le vecteur Z est :

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

Pour investir les propriétés de la statistique Z , on utilise les propriétés du processus stochastique :

$$H_n(t) = \frac{1}{\sqrt{n}} \left(N(t) - \int_0^t \lambda(u, \hat{\theta}) Y(u) du \right)$$

dans l'intervalle $[0, \tau]$, donné dans le lemme suivant.

Lemme : Soit pour $i = 1, \dots, m$; $j, j' = 1, \dots, k$

$$V_j = V(a_j) - V(a_{j-1}), \quad v_{jj'} = Cov(V_j, V_{j'}),$$

$$A_j = A(a_j) - A(a_{j-1}), \quad C_j = (C_{1j}, \dots, C_{mj})^T = C(a_j) - C(a_{j-1}),$$

$$V = [v_{jj'}]_{k \times k}, \quad C = [c_{ij'}]_{m \times k},$$

et on note par A la matrice diagonale $k \times k$, avec comme éléments diagonaux A_1, \dots, A_k .

Bagdonavičius et Nikulin (2011) ont vérifié, que sous conditions de Hjort (1990) :

$$Z \xrightarrow{d} Y \sim \mathcal{N}_k(0, V) \quad \text{quand } n \rightarrow \infty$$

où

$$V = A - C^T i^{-1}(\theta) C$$

Si G est non dégénérée alors l'inverse généralisée de la matrice V est

$$V^- = A + A^{-1} C^T G^{-1} C A^{-1}$$

où

$$G = i - C A^{-1} C^T$$

Théorème : [Bagdonavičius et al., (2010a)]

Sous conditions de Hjort (1990) les estimateurs suivants de $A_j, C_j, i(\theta)$ et V sont convergents :

$$\hat{A}_j = \frac{U_j}{n}, \quad \hat{C}_j = \frac{1}{n} \int_{a_{j-1}}^{a_j} \frac{\partial}{\partial \theta} \ln \lambda(u, \hat{\theta}) dN(t)$$

et

$$\widehat{i} = \frac{1}{n} \int_0^\tau \left(\frac{\partial}{\partial \theta} \ln \lambda(u, \widehat{\theta}) \right) \left(\frac{\partial}{\partial \theta} \ln \lambda(u, \widehat{\theta}) \right)^T dN(u), \quad \widehat{V} = \widehat{A} - \widehat{C}^T \widehat{i}^{-1} \widehat{C}$$

La matrice $\widehat{i} = -\frac{1}{n} \ddot{\ell}(\widehat{\theta})$ est aussi un estimateur convergent de $i(\theta)$ mais il est recommandé d'utiliser l'estimateur si-dessus pour assurer que les deux composantes de la statistique du test NRR sont non négatives pour chaque n

$$Y_n^2 = Z^T \widehat{V}^- Z$$

où \widehat{V}^- est l'inverse généralisée de la matrice \widehat{V} et donnée par :

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

Donc la statistique du test peut être écrite sous la forme :

$$Y_n^2 = Z^T \widehat{A}^{-1} Z + Z^T \widehat{A}^{-1} \widehat{C}^T \widehat{G}^- \widehat{C} \widehat{A}^{-1} Z = \sum_{j=1}^k \frac{(U_j - e_j)^2}{U_j} + Q$$

où

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

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

$$\widehat{i}_{ll'} = \frac{1}{n} \sum_{i=1}^n \delta_i \frac{\partial \ln \lambda(X_i; \widehat{\theta})}{\partial \theta_l} \frac{\partial \ln \lambda(X_i; \widehat{\theta})}{\partial \theta_{l'}}, \quad \widehat{C}_{lj} = \frac{1}{n} \sum_{i:X_i \in I_j}^n \delta_i \frac{\partial \ln \lambda(X_i; \widehat{\theta})}{\partial \theta_l},$$

$$\widehat{A}_j = \frac{U_j}{n}, \quad U_j = \sum_{i:X_i \in I_j}^n \delta_i, \quad Z_j = \frac{1}{\sqrt{n}} (U_j - e_j),$$

$i = 1, \dots, n, \quad j = 1, \dots, k, \quad l, l' = 1, \dots, m.$

Notons par $\widehat{g}_{ll'}$ les éléments de \widehat{G} .

La forme quadratique Q peut être écrite comme suit :

$$Q = \sum_{l=1}^m \sum_{l'=1}^m W_l g_{ll'} W_{l'}$$

Sous H_0 la distribution limite de la statistique Y_n^2 est une distribution du Chi-deux avec :

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

degrés de liberté (Si G est non dégénérée alors $r = k$).

Choix de \hat{a}_j :

Pour ce test et afin de s'assurer de l'égalité des effectifs théoriques, il est recommandé de prendre a_j comme fonctions de données aléatoires. L'idée est de diviser l'intervalle $[0, \tau]$ en k intervalles avec des nombres de défaillances prévus égaux (qui ne sont pas nécessairement des entiers). Il semble qu'il est préférable de diviser $[0, \tau]$ en intervalles avec des probabilités estimées égales selon le modèle, parce que dans la plupart des cas, les temps censurés seront concentrés sur le côté droit des données, et des petits nombres de défaillances peuvent être observés dans les intervalles d'extrémités.

Soit

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

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

Si i est le plus petit nombre naturel vérifiant :

$E_j \in [b_{i-1}, b_i]$, $j = 1, \dots, k-1$ alors :

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

d'où

$$\hat{a}_j = \Lambda^{-1} \left([E_j - \sum_{l=1}^{i-1} \Lambda(X_{(l)}, \hat{\theta})]/(n - i + 1), \hat{\theta} \right), \quad \hat{a}_k = \max(X_{(n)}, \tau),$$

où Λ^{-1} est l'inverse de la fonction de hasard cumulative Λ .

Pour ce choix d'intervalles, on a $e_j = \frac{E_k}{k}$ pour tout j .

Ainsi, en remplaçant a_j par \hat{a}_j dans l'expression de la statistique Y_n^2 , la distribution limite de la statistique Y_n^2 est encore un Chi-deux avec r degrés de liberté, comme dans le cas de a_j fixe.

3.2 Application du test NRR pour la distribution IWG

Dans cette section, nous étudions la validité du modèle inverse Weibull généralisé IWG, par un test d'ajustement basé sur la statistique NRR.

Supposons H_0 vérifiée, c'est à-dire que le taux de défaillance T_i suit une distribution inverse Weibull généralisée IWG.

La fonction de survie est :

$$S(t, \theta) = 1 - F(t, \theta) = 1 - \exp[-\gamma(\frac{\alpha}{t})^\beta]$$

Après avoir calculé les estimateurs du maximum de vraisemblance des paramètres α, β, γ , nous calculons les nombres de défaillances théoriques suivants :

$$\begin{aligned} E_j &= \sum_{i:X_i > a_{j-1}} \ln \frac{S(a_{j-1}, \hat{\theta})}{S(a_j \wedge X_i, \hat{\theta})} \\ &= \sum_{i:X_i > a_{j-1}} \ln \frac{1 - \exp[-\hat{\gamma}(\frac{\hat{\alpha}}{a_{j-1}})^{\hat{\beta}}]}{1 - \exp[-\hat{\gamma}(\frac{\hat{\alpha}}{a_j \wedge X_i})^{\hat{\beta}}]} \end{aligned}$$

Le choix de \hat{a}_j lorsque la distribution de base est une inverse Weibull généralisée, est obtenu comme suit :

$$\hat{a}_j = -\hat{\gamma}^{1/\hat{\nu}} \hat{\alpha} \left[\ln \left(1 - \exp \left[\frac{E_j - \sum_{l=1}^{i-1} \Lambda(t_l, \hat{\theta})}{n - i + 1} \right] \right) \right]^{-1/\hat{\nu}}$$

pour $j = 1, \dots, k-1$, $\hat{a}_k = t_{(n)}$,

Nous formulons la fonction risque comme suit :

$$\begin{aligned} \lambda(t, \hat{\theta}) &= -S'(t)/S(t) \\ &= \hat{\gamma} \hat{\beta} \hat{\alpha}^{\hat{\beta}} t^{-(\hat{\beta}+1)} \exp[-\hat{\gamma}(\frac{\hat{\alpha}}{t})^{\hat{\beta}}] \\ &\quad \times [1 - \exp[-\hat{\gamma}(\frac{\hat{\alpha}}{t})^{\hat{\beta}}]]^{-1}, \quad \hat{\theta} = (\hat{\alpha}, \hat{\beta}, \hat{\gamma})^T \end{aligned}$$

3.2.1 Calcul de la matrice \hat{W}

Le calcul de la matrice \hat{W} passe par celui de la matrice \hat{C} . Calculons les dérivées partielles de la fonction $\ln \lambda(t; \theta)$ par rapport à α , β et γ :

$$\frac{\partial \ln \lambda(t; \theta)}{\partial \alpha} = \frac{\hat{\beta}}{\hat{\gamma}} - \hat{\beta} \hat{\gamma} \hat{\alpha}^{\hat{\beta}-1} \left(t^{-\hat{\beta}} + \exp \left[-\hat{\gamma} \left(\frac{\hat{\alpha}}{t} \right)^{\hat{\beta}} \right] \left[1 - \exp \left[-\hat{\gamma} \left(\frac{\hat{\alpha}}{t} \right)^{\hat{\beta}} \right] \right]^{-1} \right)$$

$$\frac{\partial \ln \lambda(t; \theta)}{\partial \beta} = \frac{1}{\hat{\beta}} + \ln \left(\frac{\hat{\alpha}}{t} \right) * 1 - \hat{\gamma} \left(\frac{\hat{\alpha}}{t} \right)^{\hat{\beta}} \left(1 + \exp \left[-\hat{\gamma} \left(\frac{\hat{\alpha}}{t} \right)^{\hat{\beta}} \right] \left[1 - \exp \left[-\hat{\gamma} \left(\frac{\hat{\alpha}}{t} \right)^{\hat{\beta}} \right] \right]^{-1} \right)$$

$$\frac{\partial \ln \lambda(t; \theta)}{\partial \gamma} = \frac{1}{\hat{\gamma}} - \left(\frac{\hat{\alpha}}{t} \right)^{\hat{\beta}} \left(1 + \exp \left[-\hat{\gamma} \left(\frac{\hat{\alpha}}{t} \right)^{\hat{\beta}} \right] \left[1 - \exp \left[-\hat{\gamma} \left(\frac{\hat{\alpha}}{t} \right)^{\hat{\beta}} \right] \right]^{-1} \right)$$

Après calcul, nous trouvons les éléments de la matrice \hat{C} qui sont donnés par :

$$\begin{aligned} \hat{C}_{1j} &= \frac{1}{n} \sum_{i:x_i \in I_j}^n \delta_i \left[\frac{\hat{\beta}}{\hat{\gamma}} - \hat{\beta} \hat{\gamma} \hat{\alpha}^{\hat{\beta}-1} \left(t_i^{-\hat{\beta}} + \exp \left[-\hat{\gamma} \left(\frac{\hat{\alpha}}{t_i} \right)^{\hat{\beta}} \right] \left[1 - \exp \left[-\hat{\gamma} \left(\frac{\hat{\alpha}}{t_i} \right)^{\hat{\beta}} \right] \right]^{-1} \right) \right] \\ \hat{C}_{2j} &= \frac{1}{n} \sum_{i:x_i \in I_j}^n \delta_i \left[\frac{1}{\hat{\beta}} + \ln \left(\frac{\hat{\alpha}}{t_i} \right) * 1 - \hat{\gamma} \left(\frac{\hat{\alpha}}{t_i} \right)^{\hat{\beta}} \left(1 + h(t_i, \hat{\theta}) \right) \right] \\ \hat{C}_{3j} &= \frac{1}{n} \sum_{i:x_i \in I_j}^n \delta_i \left[\frac{1}{\hat{\gamma}} - \left(\frac{\hat{\alpha}}{t_i} \right)^{\hat{\beta}} \left(1 + \exp \left[-\hat{\gamma} \left(\frac{\hat{\alpha}}{t_i} \right)^{\hat{\beta}} \right] \left[1 - \exp \left[-\hat{\gamma} \left(\frac{\hat{\alpha}}{t_i} \right)^{\hat{\beta}} \right] \right]^{-1} \right) \right] \end{aligned}$$

3.2.2 Calcul de la matrice de Fisher estimée \hat{I}

Dans ce cas, à partir de la fonction de la log-vraisemblance, nous pouvons estimer la matrice d'information de Fisher en utilisant la formule :

$$\hat{i}_{ll'} = \frac{1}{n} \sum_{i=1}^n \delta_i \frac{\partial \ln \lambda(t_i; \hat{\theta})}{\partial \theta_l} \frac{\partial \ln \lambda(t_i; \hat{\theta})}{\partial \theta_{l'}}. \quad l, l' = 1, 2, 3.$$

Les éléments de la matrice d'information de Fisher estimée ci-dessus sont donnés comme suit :

$$\begin{aligned}
\hat{i}_{11} &= \hat{\beta}^2 \hat{\alpha}^{-2} + \frac{\hat{\beta}^2 \hat{\gamma} \hat{\alpha}^{\hat{\beta}-2}}{n} \sum_{i=1}^n \delta_i \left[t_i^{-\hat{\beta}} - \exp \left[-\hat{\gamma} \left(\frac{\hat{\alpha}}{t_i} \right)^{\hat{\beta}} \right] \left[1 - \exp \left[-\hat{\gamma} \left(\frac{\hat{\alpha}}{t_i} \right)^{\hat{\beta}} \right] \right]^{-1} \right] \\
&\quad \left[\hat{\gamma} \hat{\alpha}^{-3\hat{\beta}} \left(t_i^{-\hat{\beta}} - \exp \left[-\hat{\gamma} \left(\frac{\hat{\alpha}}{t_i} \right)^{\hat{\beta}} \right] \left[1 - \exp \left[-\hat{\gamma} \left(\frac{\hat{\alpha}}{t_i} \right)^{\hat{\beta}} \right] \right]^{-1} \right) - 2 \right] \\
\hat{i}_{22} &= \frac{1}{\hat{\beta}^2} + \frac{1}{n} \sum_{i=1}^n \delta_i \left[\ln \left(\frac{\hat{\alpha}}{t_i} \right) 1 - \hat{\gamma} \left(\frac{\hat{\alpha}}{t_i} \right)^{\hat{\beta}} \left(1 + \exp \left[-\hat{\gamma} \left(\frac{\hat{\alpha}}{t_i} \right)^{\hat{\beta}} \right] \left[1 - \exp \left[-\hat{\gamma} \left(\frac{\hat{\alpha}}{t_i} \right)^{\hat{\beta}} \right] \right]^{-1} \right) \right] \\
&\quad \left[\ln \left(\frac{\hat{\alpha}}{t_i} \right) 1 - \hat{\gamma} \left(\frac{\hat{\alpha}}{t_i} \right)^{\hat{\beta}} \left(1 + \exp \left[-\hat{\gamma} \left(\frac{\hat{\alpha}}{t_i} \right)^{\hat{\beta}} \right] \left[1 - \exp \left[-\hat{\gamma} \left(\frac{\hat{\alpha}}{t_i} \right)^{\hat{\beta}} \right] \right]^{-1} \right) + \frac{2}{\hat{\beta}} \right] \\
\hat{i}_{33} &= \frac{1}{\hat{\gamma}^2} + \frac{1}{n} \sum_{i=1}^n \delta_i \hat{\alpha}^{\hat{\beta}} t_i^{-\hat{\beta}} \left[1 + \exp \left[-\hat{\gamma} \left(\frac{\hat{\alpha}}{t_i} \right)^{\hat{\beta}} \right] \left[1 - \exp \left[-\hat{\gamma} \left(\frac{\hat{\alpha}}{t_i} \right)^{\hat{\beta}} \right] \right]^{-1} \right] \\
&\quad \left[-2 + \left(1 + \left(\frac{\hat{\alpha}}{t_i} \right)^{\hat{\beta}} \exp \left[-\hat{\gamma} \left(\frac{\hat{\alpha}}{t_i} \right)^{\hat{\beta}} \right] \left[1 - \exp \left[-\hat{\gamma} \left(\frac{\hat{\alpha}}{t_i} \right)^{\hat{\beta}} \right] \right]^{-1} \right) \right] \\
\hat{i}_{12} &= \frac{1}{n} \sum_{i=1}^n \delta_i \left[\frac{\hat{\beta}}{\hat{\gamma}} - \hat{\beta} \hat{\gamma} \hat{\alpha}^{\hat{\beta}-1} \left(t_i^{-\hat{\beta}} + \exp \left[-\hat{\gamma} \left(\frac{\hat{\alpha}}{t_i} \right)^{\hat{\beta}} \right] \left[1 - \exp \left[-\hat{\gamma} \left(\frac{\hat{\alpha}}{t_i} \right)^{\hat{\beta}} \right] \right]^{-1} \right) \right] \\
&\quad \left[\frac{1}{\hat{\beta}} + \ln \left(\frac{\hat{\alpha}}{t_i} \right) * 1 - \hat{\gamma} \left(\frac{\hat{\alpha}}{t_i} \right)^{\hat{\beta}} \left(1 + \exp \left[-\hat{\gamma} \left(\frac{\hat{\alpha}}{t_i} \right)^{\hat{\beta}} \right] \left[1 - \exp \left[-\hat{\gamma} \left(\frac{\hat{\alpha}}{t_i} \right)^{\hat{\beta}} \right] \right]^{-1} \right) \right] \\
\hat{i}_{13} &= \frac{1}{n} \sum_{i=1}^n \delta_i \left[\frac{\hat{\beta}}{\hat{\gamma}} - \hat{\beta} \hat{\gamma} \hat{\alpha}^{\hat{\beta}-1} \left(t_i^{-\hat{\beta}} + \exp \left[-\hat{\gamma} \left(\frac{\hat{\alpha}}{t_i} \right)^{\hat{\beta}} \right] \left[1 - \exp \left[-\hat{\gamma} \left(\frac{\hat{\alpha}}{t_i} \right)^{\hat{\beta}} \right] \right]^{-1} \right) \right] \\
&\quad \left[\frac{1}{\hat{\gamma}} - \left(\frac{\hat{\alpha}}{t_i} \right)^{\hat{\beta}} \left(1 + \exp \left[-\hat{\gamma} \left(\frac{\hat{\alpha}}{t_i} \right)^{\hat{\beta}} \right] \left[1 - \exp \left[-\hat{\gamma} \left(\frac{\hat{\alpha}}{t_i} \right)^{\hat{\beta}} \right] \right]^{-1} \right) \right]
\end{aligned}$$

$$\hat{i}_{23} = \frac{1}{n} \sum_{i=1}^n \delta_i \left[\frac{1}{\hat{\beta}} + \ln\left(\frac{\hat{\alpha}}{t_i}\right) * 1 - \hat{\gamma}\left(\frac{\hat{\alpha}}{t_i}\right)^{\hat{\beta}} \left(1 + \exp\left[-\hat{\gamma}\left(\frac{\hat{\alpha}}{t_i}\right)^{\hat{\beta}}\right] \left[1 - \exp\left[-\hat{\gamma}\left(\frac{\hat{\alpha}}{t_i}\right)^{\hat{\beta}}\right] \right)^{-1} \right) \right. \\ \left. \left[\frac{1}{\hat{\gamma}} - \left(\frac{\hat{\alpha}}{t_i}\right)^{\hat{\beta}} \left(1 + \exp\left[-\hat{\gamma}\left(\frac{\hat{\alpha}}{t_i}\right)^{\hat{\beta}}\right] \left[1 - \exp\left[-\hat{\gamma}\left(\frac{\hat{\alpha}}{t_i}\right)^{\hat{\beta}}\right] \right)^{-1} \right) \right]$$

Enfin, on obtient la forme explicite de la statistique Y_n^2 du test NRR

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

pour la distribution inverse Weibull généralisée IWG , comme suit :

$$Y_n^2 = \sum_{j=1}^k \frac{(U_j - e_j)^2}{U_j} + \sum_{j=1}^k \frac{1}{n} \left(\sum_{i=1}^n \delta_i \frac{\partial \ln \left(\hat{\gamma} \hat{\beta} \hat{\alpha}^{\hat{\beta}} t_i^{-(\hat{\beta}+1)} \exp[-\hat{\gamma}(\frac{\hat{\alpha}}{t_i})^{\hat{\beta}}] [1 - \exp[-\hat{\gamma}(\frac{\hat{\alpha}}{t_i})^{\hat{\beta}}]]^{-1} \right)}{\partial \theta_l} \right. \\ \left. A_j^{-1} \left(\frac{1}{\sqrt{n}} (U_j - e_j) \right)^T \left(\frac{1}{n} \sum_{i=1}^n \delta_i \frac{\partial \ln \left(\hat{\gamma} \hat{\beta} \hat{\alpha}^{\hat{\beta}} t_i^{-(\hat{\beta}+1)} \exp[-\hat{\gamma}(\frac{\hat{\alpha}}{t_i})^{\hat{\beta}}] [1 - \exp[-\hat{\gamma}(\frac{\hat{\alpha}}{t_i})^{\hat{\beta}}]]^{-1} \right)}{\partial \theta_l} \right. \right. \\ \left. \left. \frac{\partial \ln \left(\hat{\gamma} \hat{\beta} \hat{\alpha}^{\hat{\beta}} t_i^{-(\hat{\beta}+1)} \exp[-\hat{\gamma}(\frac{\hat{\alpha}}{t_i})^{\hat{\beta}}] [1 - \exp[-\hat{\gamma}(\frac{\hat{\alpha}}{t_i})^{\hat{\beta}}]]^{-1} \right)}{\partial \theta_{l'}} - \sum_{j=1}^k C_{lj} C_{l'j} A_j^{-1} \right)^{-1} \right. \\ \left. \left(\sum_{i=1}^n \delta_i \frac{\partial \ln \left(\hat{\gamma} \hat{\beta} \hat{\alpha}^{\hat{\beta}} t_i^{-(\hat{\beta}+1)} \exp[-\hat{\gamma}(\frac{\hat{\alpha}}{t_i})^{\hat{\beta}}] [1 - \exp[-\hat{\gamma}(\frac{\hat{\alpha}}{t_i})^{\hat{\beta}}]]^{-1} \right)}{\partial \theta_l} \right. \right. \\ \left. \left. A_j^{-1} \left(\frac{1}{\sqrt{n}} (U_j - e_j) \right) \right) \right)$$

où

$\hat{\theta} = (\hat{\alpha}, \hat{\beta}, \hat{\gamma})^T$, $l, l' = 1, 2, 3$, l et l' sont les indices des paramètres du vecteur $\hat{\theta}$.

3.3 Test d'ajustement pour le modèle AFT inverse Weibull généralisé en cas de données censurées

Le test d'ajustement du Chi-deux couramment utilisé dans l'analyse de la survie et de la fiabilité, est encore plus important dans les modèles à temps de vie accéléré (AFT) où un petit écart dans le choix du modèle peut changer totalement la décision finale.

Nous adaptons le test proposé par Bagdinavčius et Nikulin (2011) pour le modèle de temps de vie accéléré dont la distribution de base est une inverse Weibull généralisée (AFT-IWG).

Considérons l'hypothèse :

$$H_0 : F(x) \in \mathcal{F}_0 = \{F_0(x, \theta), x \in R^1, \theta \in \Theta \subset R^s\},$$

Où $\theta = (\theta_1, \dots, \theta_s)^T \in \Theta \subset R^s$ est le vecteur des paramètres inconnus de dimension s et F_0 est la fonction de répartition AFT- GIW. Prenons un intervalle de temps fini $[0, \tau]$ et divisons-le en $k > s$ intervalles $I_j = (a_{j-1}, a_j]$

où

$$0 = < a_0 < a_1 \dots < a_{k-1} < a_k = +\infty.$$

Dans ce cas, la valeur de \hat{a}_j estimée est donnée par :

$$\hat{a}_j = \Lambda^{-1} \left((E_j - \sum_{l=1}^{i-1} \Lambda(X_{(l)}, \hat{\theta})) / (n - i + 1), \hat{\theta} \right), \quad \hat{a}_k = X_{(n)}, \quad j = 1, \dots, k.$$

Lorsque $\hat{\theta}$ est l'estimateur du maximum de vraisemblance du paramètre θ , Λ^{-1} est l'inverse de la fonction de hasard cumulé Λ , $X_{(i)}$ et

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

Les \hat{a}_j sont des fonctions de données aléatoires telles que les k intervalles choisis ont des défaillances prévues e_j toutes égales.

3.3.1 Choix des intervalles

Dans notre cas, le choix de \hat{a}_j est obtenu comme suit :

$$\hat{a}_j = -\hat{\gamma}^{1/\hat{\nu}} \hat{\alpha} \exp \left\{ \hat{\beta}^T z_i \right\} \left[\ln \left(1 - \exp \left[\frac{E_j - \sum_{l=1}^{i-1} \Lambda(t_l, \hat{\theta})}{n - i + 1} \right] \right) \right]^{-1/\hat{\nu}}$$

pour $j = 1, \dots, k - 1$, $\hat{a}_k = t_{(n)}$,

où

$$E_j = \sum_{i:X_i>a_j} \left(\Lambda(a_j \wedge t_i, \hat{\theta}) - \Lambda(a_{j-1}, \hat{\theta}) \right)$$

et

$$E_k = \sum_{i=1}^n \Lambda(t_i, \hat{\theta})$$

Pour un tel choix d'intervalles, nous avons $e_j = E_k/k$ pour tout j .

3.3.2 Calcul de la matrice \hat{W}

Les éléments de la matrice estimée \hat{W} définie par :

$$\hat{W}_l = \sum_{j=1}^k \hat{C}_{lj} \hat{A}_j^{-1} Z_j, \quad l = 1, 2, \dots, 5. \quad j = 1, \dots, k.$$

sont obtenus de la manière suivante :

$$\begin{aligned}
\hat{C}_{1j} &= \frac{1}{n} \sum_{i:x_i \in I_j}^n \delta_i \left[\frac{\hat{\nu}}{\hat{\alpha}} - \hat{\nu} \hat{\gamma} \hat{\alpha}^{\hat{\nu}-1} (\exp [\hat{\nu} \hat{\beta}^T z_i] t_i^{-\hat{\nu}} + (\exp [\beta^T z_i] t_i^{-\nu} \right. \\
&\quad \left. \exp [-\gamma \alpha^\nu \exp [\nu \beta^T z_i] t_i^{-\nu}]]) [1 - \exp [-\gamma \alpha^\nu \exp [\nu \beta^T z_i] t_i^{-\nu}]])^{-1} \right] \\
\hat{C}_{2j} &= \frac{1}{n} \sum_{i:x_i \in I_j}^n \delta_i \left[\frac{1}{\hat{\nu}} + \ln(\hat{\alpha}) - \ln(t_i) + \hat{\beta}^T z_i - \hat{\gamma} \hat{\alpha}^{\hat{\nu}} \ln \left(\hat{\alpha} \left[\exp [\hat{\nu} \hat{\beta}^T z_i] t_i^{-\hat{\nu}} \right]^{\hat{\nu}-1} \right) \right. \\
&\quad \left. (\exp [\hat{\nu} \hat{\beta}^T z_i] t_i^{-\hat{\nu}} + [\exp [\beta^T z_i] t_i^{-\nu} \exp [-\gamma \alpha^\nu \exp [\nu \beta^T z_i] t_i^{-\nu}]] \right. \\
&\quad \left. [1 - \exp [-\gamma \alpha^\nu \exp [\nu \beta^T z_i] t_i^{-\nu}]]^{-1} \right] \\
\hat{C}_{3j} &= \frac{1}{n} \sum_{i:x_i \in I_j}^n \delta_i \left[\frac{1}{\hat{\gamma}} - \hat{\alpha}^{\hat{\nu}} (\exp [\hat{\nu} \hat{\beta}^T z_i] t_i^{-\hat{\nu}} + [\exp [\beta^T z_i] t_i^{-\nu} \right. \\
&\quad \left. \exp [-\gamma \alpha^\nu \exp [\nu \beta^T z_i] t_i^{-\nu}]] [1 - \exp [-\gamma \alpha^\nu \exp [\nu \beta^T z_i] t_i^{-\nu}]]^{-1} \right] \\
\hat{C}_{4j} &= \frac{1}{n} \sum_{i:x_i \in I_j}^n \delta_i \left[\hat{\nu} - \hat{\gamma} \hat{\alpha}^{\hat{\nu}} \hat{\nu} (\exp [\hat{\nu} \hat{\beta}^T z_i] t_i^{-\hat{\nu}} + [\exp [\beta^T z_i] t_i^{-\nu} \right. \\
&\quad \left. \exp [-\gamma \alpha^\nu \exp [\nu \beta^T z_i] t_i^{-\nu}]] [1 - \exp [-\gamma \alpha^\nu \exp [\nu \beta^T z_i] t_i^{-\nu}]]^{-1} \right] \\
\hat{C}_{5j} &= \frac{1}{n} \sum_{i:x_i \in I_j}^n \delta_i \left[\nu z_i - \hat{\gamma} \hat{\alpha}^{\hat{\nu}} \hat{\nu} z_i (\exp [\hat{\nu} \hat{\beta}^T z_i] t_i^{-\hat{\nu}} + [\exp [\beta^T z_i] t_i^{-\nu} \right. \\
&\quad \left. \exp [-\gamma \alpha^\nu \exp [\nu \beta^T z_i] t_i^{-\nu}]] [1 - \exp [-\gamma \alpha^\nu \exp [\nu \beta^T z_i] t_i^{-\nu}]]^{-1} \right]
\end{aligned}$$

Après simplifications, nous pouvons les écrire comme suit :

$$\begin{aligned}
\hat{C}_{1j} &= \frac{1}{n} \sum_{i:x_i \in I_j}^n \delta_i \left[\frac{\hat{\nu}}{\hat{\alpha}} - \hat{\nu} \hat{\gamma} \hat{\alpha}^{\hat{\nu}-1} h(t_i, \hat{\theta}) \right] \\
\hat{C}_{2j} &= \frac{1}{n} \sum_{i:x_i \in I_j}^n \delta_i \left[\frac{1}{\hat{\nu}} + \ln(\hat{\alpha}) - \ln(t_i) + \hat{\beta}^T z_i \right. \\
&\quad \left. - \hat{\gamma} \hat{\alpha}^{\hat{\nu}} \ln \left(\hat{\alpha} M^{\hat{\nu}-1}(t_i, \hat{\theta}) \right) h(t_i, \hat{\theta}) \right] \\
\hat{C}_{3j} &= \frac{1}{n} \sum_{i:x_i \in I_j}^n \delta_i \left[\frac{1}{\hat{\gamma}} - \hat{\alpha}^{\hat{\nu}} h(t_i, \hat{\theta}) \right] \\
\hat{C}_{4j} &= \frac{1}{n} \sum_{i:x_i \in I_j}^n \delta_i \left[\hat{\nu} - \hat{\gamma} \hat{\alpha}^{\hat{\nu}} \hat{\nu} h(t_i, \hat{\theta}) \right], \\
\hat{C}_{5j} &= \frac{1}{n} \sum_{i:x_i \in I_j}^n \delta_i \left[\nu z_i - \hat{\gamma} \hat{\alpha}^{\hat{\nu}} \hat{\nu} z_i h(t_i, \hat{\theta}) \right]
\end{aligned}$$

où

$$\begin{aligned}
h(t_i, \hat{\theta}) &= (\exp [\hat{\nu} \hat{\beta}^T z_i] t_i^{-\hat{\nu}} + [\exp [\beta^T z_i] t_i^{-\nu} \exp [-\gamma \alpha^\nu \exp [\nu \beta^T z_i] t_i^{-\nu}]] \\
&\quad [1 - \exp [-\gamma \alpha^\nu \exp [\nu \beta^T z_i] t_i^{-\nu}]]^{-1})
\end{aligned}$$

et

$$M^{\hat{\nu}-1}(t_i, \hat{\theta}) = \left[\exp [\hat{\nu} \hat{\beta}^T z_i] t_i^{-\hat{\nu}} \right]^{\hat{\nu}-1}$$

3.3.3 Matrice d'information de Fisher estimée \hat{I}

Les composants de la matrice d'information $\hat{I} = (\hat{i}_{ll'})_{5 \times 5}$ sont nécessaires pour le calcul de la matrice G .

Comme

$$I = -E \left(\frac{\partial^2 \ell(t, \theta)}{\partial \theta_l \partial \theta_{l'}} \right)$$

Alors nous obtenons les éléments de la matrice de Fisher estimée qui s'écrivent comme suit :

$$\begin{aligned} \frac{\partial^2 \ell(t, \theta)}{\partial \alpha \partial \alpha} &= \hat{i}_{11} = \frac{1}{n} \sum_{i=1}^n \delta_i \left(\frac{\hat{\nu}}{\hat{\alpha}} \right)^2 - \delta_i \hat{\nu}^2 \hat{\gamma} \hat{\alpha}^{\hat{\nu}-2} (\exp [\hat{\nu} \hat{\beta}^T z_i] t_i^{-\hat{\nu}} + [\exp [\beta^T z_i] t_i^{-\nu} \exp [-\gamma \alpha^\nu \exp [\nu \beta^T z_i] t_i^{-\nu}]])^{-1} \\ &\quad (2 - \hat{\gamma} \hat{\alpha}^{\hat{\nu}} (\exp [\hat{\nu} \hat{\beta}^T z_i] t_i^{-\hat{\nu}} + [\exp [\beta^T z_i] t_i^{-\nu} \exp [-\gamma \alpha^\nu \exp [\nu \beta^T z_i] t_i^{-\nu}]]) \\ &\quad [1 - \exp [-\gamma \alpha^\nu \exp [\nu \beta^T z_i] t_i^{-\nu}]]^{-1}) \\ \frac{\partial^2 \ell(t, \theta)}{\partial \alpha \partial \nu} &= \hat{i}_{12} = \frac{1}{n} \sum_{i=1}^n \delta_i \left[\frac{\hat{\nu}}{\hat{\alpha}} - \hat{\nu} \hat{\gamma} \hat{\alpha}^{\hat{\nu}-1} \exp [\hat{\nu} \hat{\beta}^T z_i] t_i^{-\hat{\nu}} \right] - \hat{\nu} \hat{\gamma} \hat{\alpha}^{\hat{\nu}-1} (\exp [\hat{\nu} \hat{\beta}^T z_i] t_i^{-\hat{\nu}} \\ &\quad + [\exp [\beta^T z_i] t_i^{-\nu} \exp [-\gamma \alpha^\nu \exp [\nu \beta^T z_i] t_i^{-\nu}]]) \\ &\quad [1 - \exp [-\gamma \alpha^\nu \exp [\nu \beta^T z_i] t_i^{-\nu}]]^{-1}) ((\frac{1}{\hat{\nu}} + \ln(\hat{\alpha}) - \ln(t) + \hat{\beta}^T z_i \\ &\quad - \hat{\gamma} \hat{\alpha}^{\hat{\nu}} \exp [\hat{\nu} \hat{\beta}^T z_i] t_i^{-\hat{\nu}} \ln (\hat{\alpha} \exp [\hat{\beta}^T z_i] t_i^{-1})) \\ &\quad + \delta_i \hat{\nu} \hat{\gamma}^2 \hat{\alpha}^{2\hat{\nu}-1} (\exp [\hat{\nu} \hat{\beta}^T z_i] t_i^{-\hat{\nu}} + [\exp [\beta^T z_i] t_i^{-\nu} \exp [-\gamma \alpha^\nu \exp [\nu \beta^T z_i] t_i^{-\nu}]]) \\ &\quad [1 - \exp [-\gamma \alpha^\nu \exp [\nu \beta^T z_i] t_i^{-\nu}]]^{-1}) \ln (\hat{\alpha} \exp [\hat{\beta}^T z_i] t_i^{-1}) \\ &\quad (\exp [\hat{\nu} \hat{\beta}^T z_i] t_i^{-\hat{\nu}} + [\exp [\beta^T z_i] t_i^{-\nu} \exp [-\gamma \alpha^\nu \exp [\nu \beta^T z_i] t_i^{-\nu}]]) \\ &\quad [1 - \exp [-\gamma \alpha^\nu \exp [\nu \beta^T z_i] t_i^{-\nu}]]^{-1}) \end{aligned}$$

$$\begin{aligned}
\frac{\partial^2 \ell(t, \theta)}{\partial \alpha \partial \gamma} = \hat{i}_{13} &= \frac{1}{n} \sum_{i=1}^n -\delta_i \hat{\nu} \hat{\alpha}^{\hat{\nu}-1} (\exp [\hat{\nu} \hat{\beta}^T z_i] t_i^{-\hat{\nu}} + [\exp [\beta^T z_i] t_i^{-\nu} \\
&\quad \exp [-\gamma \alpha^\nu \exp [\nu \beta^T z_i] t_i^{-\nu}]]) [1 - \exp [-\gamma \alpha^\nu \exp [\nu \beta^T z_i] t_i^{-\nu}]]^{-1}) \\
&\quad (2 - \hat{\gamma} \hat{\alpha}^{\hat{\nu}} (\exp [\hat{\nu} \hat{\beta}^T z_i] t_i^{-\hat{\nu}} + [\exp [\beta^T z_i] t_i^{-\nu} \exp [-\gamma \alpha^\nu \exp [\nu \beta^T z_i] t_i^{-\nu}]]) \\
&\quad [1 - \exp [-\gamma \alpha^\nu \exp [\nu \beta^T z_i] t_i^{-\nu}]]^{-1})) + \delta_i (\frac{\hat{\nu}}{\hat{\alpha} \hat{\gamma}})
\end{aligned}$$

$$\begin{aligned}
\frac{\partial^2 \ell(t, \theta)}{\partial \alpha \partial \beta_0} = \hat{i}_{14} &= \frac{1}{n} \sum_{i=1}^n -\delta_i \hat{\nu}^2 \hat{\gamma} \hat{\alpha}^{\hat{\nu}-1} (\exp [\hat{\nu} \hat{\beta}^T z_i] t_i^{-\hat{\nu}} + [\exp [\beta^T z_i] t_i^{-\nu} \\
&\quad \exp [-\gamma \alpha^\nu \exp [\nu \beta^T z_i] t_i^{-\nu}]]) [1 - \exp [-\gamma \alpha^\nu \exp [\nu \beta^T z_i] t_i^{-\nu}]]^{-1}) \\
&\quad (2 - \hat{\gamma} \hat{\alpha}^{\hat{\nu}} (\exp [\hat{\nu} \hat{\beta}^T z_i] t_i^{-\hat{\nu}} + [\exp [\beta^T z_i] t_i^{-\nu} \exp [-\gamma \alpha^\nu \exp [\nu \beta^T z_i] t_i^{-\nu}]]) \\
&\quad [1 - \exp [-\gamma \alpha^\nu \exp [\nu \beta^T z_i] t_i^{-\nu}]]^{-1})) + \delta_i [\hat{\nu}^2 \hat{\alpha}^{-1}]
\end{aligned}$$

$$\begin{aligned}
\frac{\partial^2 \ell(t, \theta)}{\partial \alpha \partial \beta_1} = \hat{i}_{15} &= \frac{1}{n} \sum_{i=1}^n -\delta_i \hat{\nu}^2 \hat{\gamma} \hat{\alpha}^{\hat{\nu}-1} z (\exp [\hat{\nu} \hat{\beta}^T z_i] t_i^{-\hat{\nu}} + [\exp [\beta^T z_i] t_i^{-\nu} \\
&\quad \exp [-\gamma \alpha^\nu \exp [\nu \beta^T z_i] t_i^{-\nu}]]) [1 - \exp [-\gamma \alpha^\nu \exp [\nu \beta^T z_i] t_i^{-\nu}]]^{-1}) \\
&\quad (2 - \hat{\gamma} \hat{\alpha}^{\hat{\nu}} (\exp [\hat{\nu} \hat{\beta}^T z_i] t_i^{-\hat{\nu}} + [\exp [\beta^T z_i] t_i^{-\nu} \exp [-\gamma \alpha^\nu \exp [\nu \beta^T z_i] t_i^{-\nu}]]) \\
&\quad [1 - \exp [-\gamma \alpha^\nu \exp [\nu \beta^T z_i] t_i^{-\nu}]]^{-1})) + \delta_i [\hat{\nu}^2 z \hat{\alpha}^{-1}]
\end{aligned}$$

$$\begin{aligned}
\frac{\partial^2 \ell(t, \theta)}{\partial \nu \partial \nu} = \hat{i}_{22} &= \frac{1}{n} \sum_{i=1}^n \delta_i (\left[\frac{1}{\hat{\nu}} + \ln(\hat{\alpha}) - \ln(t) + \hat{\beta}^T z_i \right] \\
&\quad - \delta_i \hat{\gamma} \hat{\alpha}^{\hat{\nu}} \ln \left(\hat{\alpha} \left[\exp [\hat{\nu} \hat{\beta}^T z_i] t_i^{-\hat{\nu}} \right]^{\hat{\nu}-1} \right) (\exp [\hat{\nu} \hat{\beta}^T z_i] t_i^{-\hat{\nu}} + [\exp [\beta^T z_i] t_i^{-\nu} \\
&\quad \exp [-\gamma \alpha^\nu \exp [\nu \beta^T z_i] t_i^{-\nu}]) [1 - \exp [-\gamma \alpha^\nu \exp [\nu \beta^T z_i] t_i^{-\nu}]]^{-1})^2
\end{aligned}$$

$$\frac{\partial^2 \ell(t, \theta)}{\partial \nu \partial \gamma} = \hat{i}_{23} = \frac{1}{n} \sum_{i=1}^n \delta_i \left(\frac{1}{\hat{\nu}} + \ln(\hat{\alpha}) - \ln(t) + \hat{\beta}^T z_i - \hat{\gamma} \hat{\alpha}^{\hat{\nu}} \ln \left(\hat{\alpha} \left[\exp \left[\hat{\nu} \hat{\beta}^T z_i \right] t_i^{-\hat{\nu}} \right]^{\hat{\nu}-1} \right) \right.$$

$$\begin{aligned} & (\exp \left[\hat{\nu} \hat{\beta}^T z_i \right] t_i^{-\hat{\nu}} + [\exp \left[\beta^T z_i \right] t_i^{-\nu} \exp \left[-\gamma \alpha^\nu \exp \left[\nu \beta^T z_i \right] t_i^{-\nu} \right]] \\ & [1 - \exp \left[-\gamma \alpha^\nu \exp \left[\nu \beta^T z_i \right] t_i^{-\nu} \right]]^{-1}) [\frac{1}{\hat{\gamma}} - \hat{\alpha}^{\hat{\nu}} \exp \left[\hat{\nu} \hat{\beta}^T z_i \right] t_i^{-\hat{\nu}} \\ & - \hat{\alpha}^{\hat{\nu}} (\exp \left[\beta^T z_i \right] t_i^{-\nu} \exp \left[-\gamma \alpha^\nu \exp \left[\nu \beta^T z_i \right] t_i^{-\nu} \right]] \\ & [1 - \exp \left[-\gamma \alpha^\nu \exp \left[\nu \beta^T z_i \right] t_i^{-\nu} \right]]^{-1})] \end{aligned}$$

$$\frac{\partial^2 \ell(t, \theta)}{\partial \nu \partial \beta_0} = \hat{i}_{24} = \frac{1}{n} \sum_{i=1}^n \delta_i \left(\frac{1}{\hat{\nu}} + \ln(\hat{\alpha}) - \ln(t) + \hat{\beta}^T z_i - \hat{\gamma} \hat{\alpha}^{\hat{\nu}} \ln \left(\hat{\alpha} \left[\exp \left[\hat{\nu} \hat{\beta}^T z_i \right] t_i^{-\hat{\nu}} \right]^{\hat{\nu}-1} \right) \right)$$

$$\begin{aligned} & (\exp \left[\hat{\nu} \hat{\beta}^T z_i \right] t_i^{-\hat{\nu}} + [\exp \left[\beta^T z_i \right] t_i^{-\nu} \exp \left[-\gamma \alpha^\nu \exp \left[\nu \beta^T z_i \right] t_i^{-\nu} \right]] \\ & [1 - \exp \left[-\gamma \alpha^\nu \exp \left[\nu \beta^T z_i \right] t_i^{-\nu} \right]]^{-1}) [\hat{\nu} z_i - \hat{\gamma} \hat{\alpha}^{\hat{\nu}} \hat{\nu} [\exp \left[\hat{\nu} \hat{\beta}^T z_i \right] t_i^{-\hat{\nu}} \\ & + ([\exp \left[\beta^T z_i \right] t_i^{-\nu} \exp \left[-\gamma \alpha^\nu \exp \left[\nu \beta^T z_i \right] t_i^{-\nu} \right]] \\ & [1 - \exp \left[-\gamma \alpha^\nu \exp \left[\nu \beta^T z_i \right] t_i^{-\nu} \right]]^{-1})]] \end{aligned}$$

$$\frac{\partial^2 \ell(t, \theta)}{\partial \nu \partial \beta_1} = \hat{i}_{25} = \frac{1}{n} \sum_{i=1}^n \delta_i \left(\frac{1}{\hat{\nu}} + \ln(\hat{\alpha}) - \ln(t) + \hat{\beta}^T z_i - \hat{\gamma} \hat{\alpha}^{\hat{\nu}} \ln \left(\hat{\alpha} \left[\exp \left[\hat{\nu} \hat{\beta}^T z_i \right] t_i^{-\hat{\nu}} \right]^{\hat{\nu}-1} \right) \right)$$

$$\begin{aligned} & (\exp \left[\hat{\nu} \hat{\beta}^T z_i \right] t_i^{-\hat{\nu}} + [\exp \left[\beta^T z_i \right] t_i^{-\nu} \exp \left[-\gamma \alpha^\nu \exp \left[\nu \beta^T z_i \right] t_i^{-\nu} \right]] \\ & [1 - \exp \left[-\gamma \alpha^\nu \exp \left[\nu \beta^T z_i \right] t_i^{-\nu} \right]]^{-1}, \hat{\theta}) [\hat{\nu} z_i - \hat{\gamma} \hat{\alpha}^{\hat{\nu}} \hat{\nu} z_i [\exp \left[\hat{\nu} \hat{\beta}^T z_i \right] t_i^{-\hat{\nu}} \\ & + ([\exp \left[\beta^T z_i \right] t_i^{-\nu} \exp \left[-\gamma \alpha^\nu \exp \left[\nu \beta^T z_i \right] t_i^{-\nu} \right]] \\ & [1 - \exp \left[-\gamma \alpha^\nu \exp \left[\nu \beta^T z_i \right] t_i^{-\nu} \right]]^{-1})]] \end{aligned}$$

$$\begin{aligned}
\frac{\partial^2 \ell(t, \theta)}{\partial \gamma \partial \gamma} &= \hat{i}_{33} = \frac{1}{n} \sum_{i=1}^n \delta_i [\hat{\gamma}^{-2} - \hat{\alpha}^{\hat{\nu}} (\exp [\hat{\nu} \hat{\beta}^T z_i] t_i^{-\hat{\nu}} + [\exp [\beta^T z_i] t_i^{-\nu} \\
&\quad \exp [-\gamma \alpha^\nu \exp [\nu \beta^T z_i] t_i^{-\nu}]]) [1 - \exp [-\gamma \alpha^\nu \exp [\nu \beta^T z_i] t_i^{-\nu}]]^{-1}) \\
&\quad (2 \hat{\gamma}^{-1} - \hat{\alpha}^{\hat{\nu}} (\exp [\hat{\nu} \hat{\beta}^T z_i] t_i^{-\hat{\nu}} + [\exp [\beta^T z_i] t_i^{-\nu} \exp [-\gamma \alpha^\nu \exp [\nu \beta^T z_i] t_i^{-\nu}]] \\
&\quad [1 - \exp [-\gamma \alpha^\nu \exp [\nu \beta^T z_i] t_i^{-\nu}]]^{-1}))] \\
\\
\frac{\partial^2 \ell(t, \theta)}{\partial \gamma \partial \beta_0} &= \hat{i}_{34} = \frac{1}{n} \sum_{i=1}^n \delta_i [\hat{\gamma}^{-1} \hat{\nu} - \hat{\nu} \hat{\alpha}^{\hat{\nu}} (\exp [\hat{\nu} \hat{\beta}^T z_i] t_i^{-\hat{\nu}} + [\exp [\beta^T z_i] t_i^{-\nu} \\
&\quad \exp [-\gamma \alpha^\nu \exp [\nu \beta^T z_i] t_i^{-\nu}]]) [1 - \exp [-\gamma \alpha^\nu \exp [\nu \beta^T z_i] t_i^{-\nu}]]^{-1}) \\
&\quad (2 - \hat{\gamma} \hat{\alpha}^{\hat{\nu}} (\exp [\hat{\nu} \hat{\beta}^T z_i] t_i^{-\hat{\nu}} + [\exp [\beta^T z_i] t_i^{-\nu} \exp [-\gamma \alpha^\nu \exp [\nu \beta^T z_i] t_i^{-\nu}]] \\
&\quad [1 - \exp [-\gamma \alpha^\nu \exp [\nu \beta^T z_i] t_i^{-\nu}]]^{-1}))] \\
\\
\frac{\partial^2 \ell(t, \theta)}{\partial \nu \partial \beta_1} &= \hat{i}_{35} = \frac{1}{n} \sum_{i=1}^n \delta_i [\hat{\gamma}^{-1} \hat{\nu} z - \hat{\nu} \hat{\alpha}^{\hat{\nu}} z (\exp [\hat{\nu} \hat{\beta}^T z_i] t_i^{-\hat{\nu}} + [\exp [\beta^T z_i] t_i^{-\nu} \\
&\quad \exp [-\gamma \alpha^\nu \exp [\nu \beta^T z_i] t_i^{-\nu}]]) [1 - \exp [-\gamma \alpha^\nu \exp [\nu \beta^T z_i] t_i^{-\nu}]]^{-1}) \\
&\quad (2 - \hat{\gamma} \hat{\alpha}^{\hat{\nu}} (\exp [\hat{\nu} \hat{\beta}^T z_i] t_i^{-\hat{\nu}} + [\exp [\beta^T z_i] t_i^{-\nu} \exp [-\gamma \alpha^\nu \exp [\nu \beta^T z_i] t_i^{-\nu}]] \\
&\quad [1 - \exp [-\gamma \alpha^\nu \exp [\nu \beta^T z_i] t_i^{-\nu}]]^{-1}))] \\
\\
\frac{\partial^2 \ell(t, \theta)}{\partial \beta_0 \partial \beta_0} &= \hat{i}_{44} = \frac{1}{n} \sum_{i=1}^n \delta_i [\hat{\nu}^2 - \hat{\gamma} \hat{\alpha}^{\hat{\nu}} \hat{\nu}^2 z^2 (\exp [\hat{\nu} \hat{\beta}^T z_i] t_i^{-\hat{\nu}} + [\exp [\beta^T z_i] t_i^{-\nu} \\
&\quad \exp [-\gamma \alpha^\nu \exp [\nu \beta^T z_i] t_i^{-\nu}]]) [1 - \exp [-\gamma \alpha^\nu \exp [\nu \beta^T z_i] t_i^{-\nu}]]^{-1}) \\
&\quad (2 - \hat{\gamma} \hat{\alpha}^{\hat{\nu}} (\exp [\hat{\nu} \hat{\beta}^T z_i] t_i^{-\hat{\nu}} + [\exp [\beta^T z_i] t_i^{-\nu} \exp [-\gamma \alpha^\nu \exp [\nu \beta^T z_i] t_i^{-\nu}]] \\
&\quad [1 - \exp [-\gamma \alpha^\nu \exp [\nu \beta^T z_i] t_i^{-\nu}]]^{-1}))]
\end{aligned}$$

$$\begin{aligned}
\frac{\partial^2 \ell(t, \theta)}{\partial \beta_0 \partial \beta_1} &= \hat{i}_{45} = \frac{1}{n} \sum_{i=1}^n \delta_i [\hat{\nu}^2 z - \hat{\nu}^2 \hat{\gamma} \hat{\alpha}^{\hat{\nu}} z (\exp [\hat{\nu} \hat{\beta}^T z_i] t_i^{-\hat{\nu}} + [\exp [\beta^T z_i] t_i^{-\nu} \\
&\quad \exp [-\gamma \alpha^\nu \exp [\nu \beta^T z_i] t_i^{-\nu}]])^{-1}) \\
&\quad (2 - \hat{\gamma} \hat{\alpha}^{\hat{\nu}} (\exp [\hat{\nu} \hat{\beta}^T z_i] t_i^{-\hat{\nu}} + [\exp [\beta^T z_i] t_i^{-\nu} \exp [-\gamma \alpha^\nu \exp [\nu \beta^T z_i] t_i^{-\nu}]]) \\
&\quad [1 - \exp [-\gamma \alpha^\nu \exp [\nu \beta^T z_i] t_i^{-\nu}]])^{-1})] \\
\frac{\partial^2 \ell(t, \theta)}{\partial \beta_1 \partial \beta_1} &= \hat{i}_{55} = \frac{1}{n} \sum_{i=1}^n \delta_i [\hat{\nu}^2 z^2 - \hat{\gamma} \hat{\alpha}^{\hat{\nu}} \hat{\nu}^2 z^2 (\exp [\hat{\nu} \hat{\beta}^T z_i] t_i^{-\hat{\nu}} + [\exp [\beta^T z_i] t_i^{-\nu} \\
&\quad \exp [-\gamma \alpha^\nu \exp [\nu \beta^T z_i] t_i^{-\nu}]])^{-1}) \\
&\quad (2 - \hat{\gamma} \hat{\alpha}^{\hat{\nu}} (\exp [\hat{\nu} \hat{\beta}^T z_i] t_i^{-\hat{\nu}} + [\exp [\beta^T z_i] t_i^{-\nu} \exp [-\gamma \alpha^\nu \exp [\nu \beta^T z_i] t_i^{-\nu}]]) \\
&\quad [1 - \exp [-\gamma \alpha^\nu \exp [\nu \beta^T z_i] t_i^{-\nu}]])^{-1})]
\end{aligned}$$

Après simplifications, on obtient les éléments de la matrice d'information de Fisher estimée :

$$\begin{aligned}
\hat{i}_{11} &= \frac{1}{n} \sum_{i=1}^n \delta_i \left[\left(\frac{\hat{\nu}}{\hat{\alpha}} \right)^2 - \hat{\nu}^2 \hat{\gamma} \hat{\alpha}^{\hat{\nu}-2} h(t_i, \hat{\theta}) \left(2 - \hat{\gamma} \hat{\alpha}^{\hat{\nu}} h(t_i, \hat{\theta}) \right) \right], \\
\hat{i}_{12} &= \frac{1}{n} \sum_{i=1}^n \delta_i \left[\frac{\hat{\nu}}{\hat{\alpha}} - \hat{\nu} \hat{\gamma} \hat{\alpha}^{\hat{\nu}-1} \exp [\hat{\nu} \hat{\beta}^T z_i] t_i^{-\hat{\nu}} - \hat{\nu} \hat{\gamma} \hat{\alpha}^{\hat{\nu}-1} W(t_i, \hat{\theta}) \right] \left[\frac{1}{\hat{\nu}} + \ln(\hat{\alpha}) \right. \\
&\quad \left. - \ln(t_i) + \hat{\beta}^T z_i - \hat{\gamma} \hat{\alpha}^{\hat{\nu}} \exp [\hat{\nu} \hat{\beta}^T z_i] t_i^{-\hat{\nu}} \ln \left(\hat{\alpha} \exp [\hat{\beta}^T z_i] t_i^{-1} \right) \right. \\
&\quad \left. - \hat{\gamma} \hat{\alpha}^{\hat{\nu}} \ln \left(\hat{\alpha} \exp [\hat{\beta}^T z_i] t_i^{-1} \right) W(t_i, \hat{\theta}) \right], \\
\hat{i}_{13} &= \frac{1}{n} \sum_{i=1}^n \delta_i \left[\frac{\hat{\nu}}{\hat{\alpha} \hat{\gamma}} - \hat{\nu} \hat{\alpha}^{\hat{\nu}-1} h(t_i, \hat{\theta}) \left(2 - \hat{\gamma} \hat{\alpha}^{\hat{\nu}} h(t_i, \hat{\theta}) \right) \right] \\
\hat{i}_{14} &= \frac{1}{n} \sum_{i=1}^n \delta_i \left[\hat{\nu}^2 \hat{\alpha}^{-1} - \hat{\nu}^2 \hat{\gamma} \hat{\alpha}^{\hat{\nu}-1} h(t_i, \hat{\theta}) \left(2 - \hat{\gamma} \hat{\alpha}^{\hat{\nu}} h(t_i, \hat{\theta}) \right) \right]
\end{aligned}$$

$$\begin{aligned}
\hat{i}_{15} &= \frac{1}{n} \sum_{i=1}^n \delta_i \left[\hat{\nu}^2 z \hat{\alpha}^{-1} - \hat{\nu}^2 \hat{\gamma} \hat{\alpha}^{\hat{\nu}-1} z h(t_i, \hat{\theta}) \left(2 - \hat{\gamma} \hat{\alpha}^{\hat{\nu}} h(t_i, \hat{\theta}) \right) \right] \\
\hat{i}_{22} &= \frac{1}{n} \sum_{i=1}^n \delta_i \left[\frac{1}{\hat{\nu}} + \ln(\hat{\alpha}) - \ln(t_i) + \hat{\beta}^T z_i - \hat{\gamma} \hat{\alpha}^{\hat{\nu}} \ln \left(\hat{\alpha} M^{\hat{\nu}-1}(t_i, \hat{\theta}) \right) h(t_i, \hat{\theta}) \right]^2 \\
\hat{i}_{23} &= \frac{1}{n} \sum_{i=1}^n \delta_i \left[\frac{1}{\hat{\nu}} + \ln(\hat{\alpha}) - \ln(t_i) + \hat{\beta}^T z_i - \hat{\gamma} \hat{\alpha}^{\hat{\nu}} \ln \left(\hat{\alpha} M^{\hat{\nu}-1}(t_i, \hat{\theta}) \right) h(t_i, \hat{\theta}) \right] \\
&\quad \left[\frac{1}{\hat{\gamma}} - \hat{\alpha}^{\hat{\nu}} \exp \left[\hat{\nu} \hat{\beta}^T z_i \right] t_i^{-\hat{\nu}} - \hat{\alpha}^{\hat{\nu}} W(t_i, \hat{\theta}) \right] \\
\hat{i}_{24} &= \frac{1}{n} \sum_{i=1}^n \delta_i \left[\frac{1}{\hat{\nu}} + \ln(\hat{\alpha}) - \ln(t_i) + \hat{\beta}^T z_i - \hat{\gamma} \hat{\alpha}^{\hat{\nu}} \ln \left(\hat{\alpha} M^{\hat{\nu}-1}(t_i, \hat{\theta}) \right) h(t_i, \hat{\theta}) \right] \\
&\quad \left[\hat{\nu} z - \hat{\gamma} \hat{\alpha}^{\hat{\nu}} \hat{\nu} \left[\exp \left[\hat{\nu} \hat{\beta}^T z_i \right] t_i^{-\hat{\nu}} + W(t_i, \hat{\theta}) \right] \right] \\
\hat{i}_{25} &= \frac{1}{n} \sum_{i=1}^n \delta_i \left[\frac{1}{\hat{\nu}} + \ln(\hat{\alpha}) - \ln(t_i) + \hat{\beta}^T z_i - \hat{\gamma} \hat{\alpha}^{\hat{\nu}} \ln \left(\hat{\alpha} M^{\hat{\nu}-1}(t_i, \hat{\theta}) \right) h(t_i, \hat{\theta}) \right] \\
&\quad \left[\hat{\nu} z_i - \hat{\gamma} \hat{\alpha}^{\hat{\nu}} \hat{\nu} z_i \left[\exp \left[\hat{\nu} \hat{\beta}^T z_i \right] t_i^{-\hat{\nu}} + W(t_i, \hat{\theta}) \right] \right] \\
\hat{i}_{33} &= \frac{1}{n} \sum_{i=1}^n \delta_i \left[\hat{\gamma}^{-2} - \hat{\alpha}^{\hat{\nu}} h(t_i, \hat{\theta}) \left(2 \hat{\gamma}^{-1} - \hat{\alpha}^{\hat{\nu}} h(t_i, \hat{\theta}) \right) \right] \\
\hat{i}_{34} &= \frac{1}{n} \sum_{i=1}^n \delta_i \left[\hat{\gamma}^{-1} \hat{\nu} - \hat{\nu} \hat{\alpha}^{\hat{\nu}} h(t_i, \hat{\theta}) \left(2 - \hat{\gamma} \hat{\alpha}^{\hat{\nu}} h(t_i, \hat{\theta}) \right) \right] \\
\hat{i}_{35} &= \frac{1}{n} \sum_{i=1}^n \delta_i \left[\hat{\gamma}^{-1} \hat{\nu} z - \hat{\nu} \hat{\alpha}^{\hat{\nu}} z h(t_i, \hat{\theta}) \left(2 - \hat{\gamma} \hat{\alpha}^{\hat{\nu}} h(t_i, \hat{\theta}) \right) \right]
\end{aligned}$$

$$\hat{i}_{44} = \frac{1}{n} \sum_{i=1}^n \left[\hat{\nu}^2 - \hat{\gamma} \hat{\alpha}^{\hat{\nu}} \hat{\nu}^2 z^2 h(t_i, \hat{\theta}) \left[2 - \hat{\gamma} \hat{\alpha}^{\hat{\nu}} h(t_i, \hat{\theta}) \right] \right]$$

$$\hat{i}_{45} = \frac{1}{n} \sum_{i=1}^n \delta_i [\hat{\nu}^2 z - \hat{\nu}^2 \hat{\gamma} \hat{\alpha}^{\hat{\nu}} z h(t_i, \hat{\theta}) \left(2 - \hat{\gamma} \hat{\alpha}^{\hat{\nu}} h(t_i, \hat{\theta}) \right)]$$

$$\hat{i}_{55} = \frac{1}{n} \sum_{i=1}^n \delta_i \left[\hat{\nu}^2 z^2 - \hat{\gamma} \hat{\alpha}^{\hat{\nu}} \hat{\nu}^2 z^2 h(t_i, \hat{\theta}) \left[2 - \hat{\gamma} \hat{\alpha}^{\hat{\nu}} h(t_i, \hat{\theta}) \right] \right]$$

Où

$$h(t_i, \hat{\theta}) = (\exp \left[\hat{\nu} \hat{\beta}^T z_i \right] t_i^{-\hat{\nu}} + [\exp \left[\beta^T z_i \right] t_i^{-\nu} \exp \left[-\gamma \alpha^\nu \exp \left[\nu \beta^T z_i \right] t_i^{-\nu} \right]] \\ [1 - \exp \left[-\gamma \alpha^\nu \exp \left[\nu \beta^T z_i \right] t_i^{-\nu} \right]]^{-1})$$

$$M^{\hat{\nu}-1}(t_i, \hat{\theta}) = \left[\exp \left[\hat{\nu} \hat{\beta}^T z_i \right] t_i^{-\hat{\nu}} \right]^{\hat{\nu}-1}$$

et

$$W(t_i, \hat{\theta}) = ([\exp \left[\beta^T z_i \right] t_i^{-\nu} \exp \left[-\gamma \alpha^\nu \exp \left[\nu \beta^T z_i \right] t_i^{-\nu} \right]] \\ [1 - \exp \left[-\gamma \alpha^\nu \exp \left[\nu \beta^T z_i \right] t_i^{-\nu} \right]]^{-1})$$

Ensuite, nous obtenons la statistique Y_n^2 :

$$Y^2 = \sum_{j=1}^k \frac{(U_j - e_j)^2}{U_j} + \hat{W}^T \times \left[\hat{i}_{ll'} - \sum_{j=1}^k \hat{C}_{lj} \hat{C}_{l'j} \hat{A}_j^{-1} \right]^- \times \hat{W}, \quad l, l' = 1, 2, \dots, 5.$$

La statistique Y_n^2 pour le modèle AFT-IWG s'écrit alors :

$$\begin{aligned}
Y_n^2 &= \sum_{j=1}^k \frac{(U_j - e_j)^2}{U_j} + \sum_{j=1}^k \frac{1}{n} \left(\sum_{i=1}^n \delta_i \frac{\partial \ln \left(\gamma \nu \alpha^\nu \exp[\nu \beta^T z] t_i^{-(\nu+1)} \exp[-\gamma \alpha^\nu \exp \{ \nu \beta^T z \} t_i^{-\nu}] \right)}{\partial \theta_l} \right. \\
&\quad \left. \frac{\partial \ln ([1 - \exp[-\gamma \alpha^\nu \exp \{ \nu \beta^T z \} t_i^{-\nu}]]^{-1})}{\partial \theta_l} A_j^{-1} \left(\frac{1}{\sqrt{n}} (U_j - e_j) \right)^T \right) \\
&\quad \left(\frac{1}{n} \sum_{i=1}^n \delta_i \frac{\partial \ln \left(\gamma \nu \alpha^\nu \exp[\nu \beta^T z] t_i^{-(\nu+1)} \exp[-\gamma \alpha^\nu \exp \{ \nu \beta^T z \} t_i^{-\nu}] \right)}{\partial \theta_l} \right. \\
&\quad \left. \frac{\partial \ln ([1 - \exp[-\gamma \alpha^\nu \exp \{ \nu \beta^T z \} t_i^{-\nu}]]^{-1})}{\partial \theta_l} \right. \\
&\quad \left. \frac{\partial \ln \left(\gamma \nu \alpha^\nu \exp[\nu \beta^T z] t_i^{-(\nu+1)} \exp[-\gamma \alpha^\nu \exp \{ \nu \beta^T z \} t_i^{-\nu}] \right)}{\partial \theta_{l'}} \right. \\
&\quad \left. \frac{\partial \ln ([1 - \exp[-\gamma \alpha^\nu \exp \{ \nu \beta^T z \} t_i^{-\nu}]]^{-1})}{\partial \theta_{l'}} - \sum_{j=1}^k C_{lj} C_{l'j} A_j^{-1})^{-1} \right. \\
&\quad \left. \left(\sum_{i=1}^n \delta_i \frac{\partial \ln \left(\gamma \nu \alpha^\nu \exp[\nu \beta^T z] t_i^{-(\nu+1)} \exp[-\gamma \alpha^\nu \exp \{ \nu \beta^T z \} t_i^{-\nu}] \right)}{\partial \theta_l} \right. \right. \\
&\quad \left. \left. \frac{\partial \ln ([1 - \exp[-\gamma \alpha^\nu \exp \{ \nu \beta^T z \} t_i^{-\nu}]]^{-1})}{\partial \theta_{l'}} A_j^{-1} \left(\frac{1}{\sqrt{n}} (U_j - e_j) \right) \right) \right)
\end{aligned}$$

où

$$\hat{\theta} = (\hat{\alpha}, \hat{\nu}, \hat{\gamma}, \hat{\beta}_0, \hat{\beta}_1)^T \quad \text{et} \quad l, l' = 1, 2, \dots, 5.$$

La distribution limite de la statistique Y_n^2 est un Chi-deux avec k degrés de liberté.

Chapitre 4

Simulations et applications numériques

4.1 Simulation de la distribution IWG

4.1.1 Estimateurs du maximum de vraisemblance des paramètres de la distribution IWG

Supposons que le modèle IWG est considéré. Les données ont été simulées $N = 10,000$ fois, avec les valeurs des paramètres $\alpha = 2$, $\beta = 3$ et $\gamma = 2$. On utilise le logiciel R et l'algorithme Barzilai-Borwein (BB) (Ravi. V 2009) pour les calculs.

Les valeurs de la moyenne des estimateurs du maximum de vraisemblance simulés $\hat{\alpha}$, $\hat{\beta}$ et $\hat{\gamma}$, et leurs **erreurs quadratiques moyennes** (notées MSE) sont calculées et données dans le tableau 4.1, (tailles des échantillons $n = 150$, $n = 350$ et $n = 550$).

$N = 10,000$	$n = 150$	$n = 350$	$n = 550$
$\hat{\alpha}$	2.1273	2.1515	2.2953
MSE	0.0204	0.0095	0.1292
$\hat{\beta}$	2.9577	3.1047	3.1155
MSE	0.0365	0.0251	0.0407
$\hat{\gamma}$	2.2125	2.3987	2.4952
MSE	0.0140	0.0126	0.0101

Tab 4.1 Erreur quadratique moyenne des EMV des paramètres de l'IWG

Du tableau 4.1, nous pouvons remarquer que les erreurs quadratiques moyennes sont en général < 0.01 .

Pour étudier **la convergence des estimateurs** du maximum de vraisemblance des paramètres α , β et γ de la distribution inverse Weibull généralisée IWG, on calcule l'erreur absolue moyenne des paramètres estimés par la méthode du maximum de vraisemblance par rapport à leurs valeurs réelles en fonction de la taille n des échantillons.

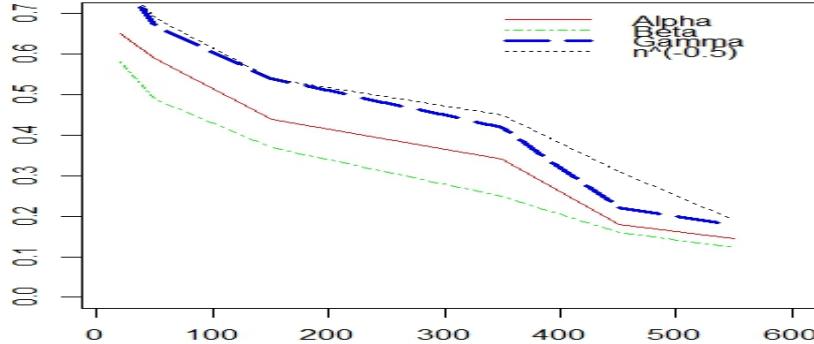


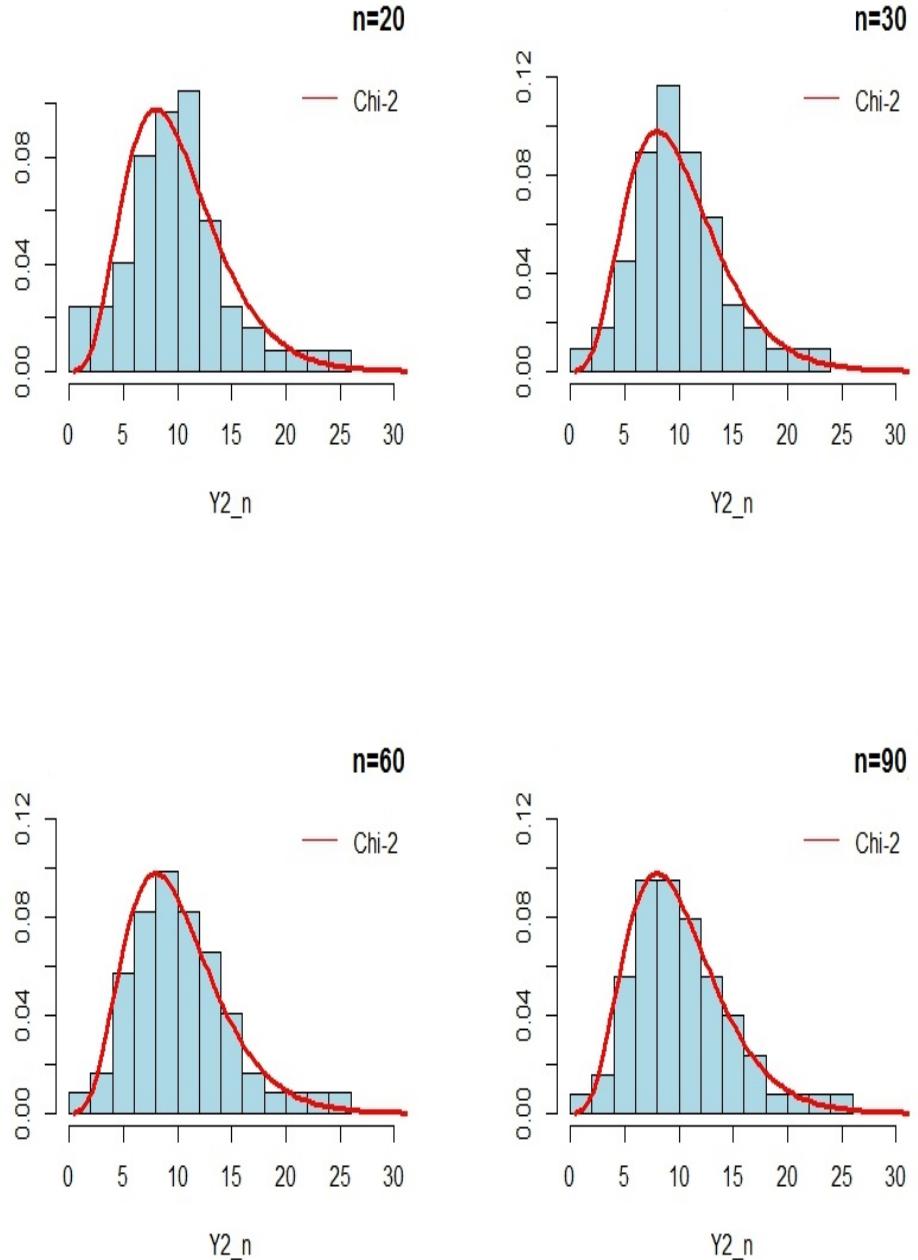
Fig. 4.1. Erreur absolue moyenne pour EMV $\hat{\alpha}$, $\hat{\beta}$ et $\hat{\gamma}$ par rapport à leurs valeurs réelles en fonction de la taille de l'échantillon n .

A partir de ces résultats (Fig. 4.1), on peut voir que tous les estimateurs convergent plus vite que $n^{-0,5}$, ce qui confirme le fait bien connu que **les estimateurs du maximum de vraisemblance sont \sqrt{n} -convergents**.

4.1.2 Distribution de la statistique Y_n^2 pour la distribution IWG

Pour étudier empiriquement **le comportement de la statistique** (NRR) Y_n^2 pour l'IWG, nous avons généré à la base de la distribution inverse Weibull généralisée, des échantillons de taille $n = 20, 30, 60, 90, 120, 250$ et 350 , et pour diverses combinaisons des paramètres α , β et γ , avec un seuil $\alpha = 0.05$.

D'autre part, nous calculons pour $N = 10,000$ fois, la distribution simulée de la statistique Y_n^2 sous la distribution inverse Weibull généralisée IWG, avec des valeurs différentes des paramètres α , β et γ , et des nombres différents de classes k . Nous représentons dans la figure 4.2 les histogrammes des statistiques Y_n^2 calculées précédemment, et la distribution Chi-deux avec k degrés de liberté correspondante.



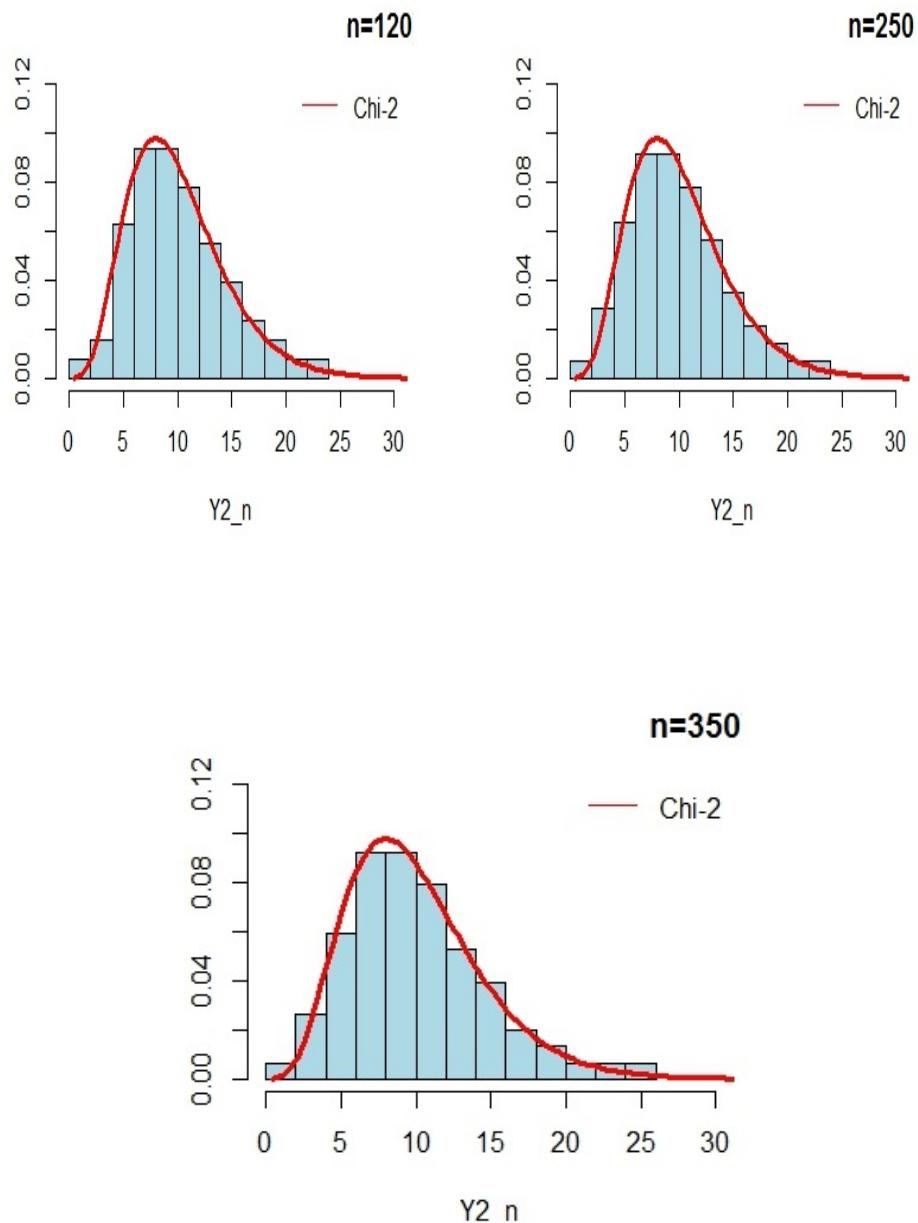


Fig. 4.2 Histogrammes de Y_n^2 pour $(\alpha = 2, \beta = 3, \gamma = 2)$, et $n = 20, 30, 60, 90, 120, 250, 350$.

Comme prévu, la statistique Y_n^2 du test NRR pour la distribution IWG, suit à la limite, la distribution du Chi-deux avec k degrés de liberté. Donc, nous concluons que **la statistique Y_n^2 est invariante**.

4.1.3 Test Y_n^2

Nous varions les valeurs des paramètres de la distribution inverse Weibull généralisée (α, β, γ) , et nous simulons $N = 10,000$ fois des échantillons de différentes tailles $n = 20, 30, 60, 90, 120, 250, 350$ et 450 .

Calcul des Niveaux empiriques N.E.

Pour chaque échantillon simulé, nous calculons la statistique Y_n^2 du test NRR. Puis, nous calculons le niveau empirique N.E qui compte le nombre de fois où $Y_n^2 \leq C_\alpha = \chi_{k,1-\alpha}^2$ divisé par $N = 10,000$.

Les tableaux 4.2 à 4.5 représentent **le cas où le paramètre α varie**.

$(\alpha, \beta, \gamma) = (0.5, 0.5, 0.5)$									
n	20	30	60	90	100	120	250	350	450
$N.E.$	0.9520	0.9540	0.9486	0.9532	0.9516	0.9516	0.9418	0.9558	0.9512
$(\alpha, \beta, \gamma) = (1.5, 0.5, 0.5)$									
n	20	30	60	90	100	120	250	350	450
$N.E.$	0.9532	0.9534	0.9520	0.9448	0.9454	0.9474	0.9482	0.9520	0.9512
$(\alpha, \beta, \gamma) = (3, 0.5, 0.5)$									
n	20	30	60	90	100	120	250	350	450
$N.E.$	0.9558	0.9562	0.9540	0.9500	0.9552	0.9518	0.9562	0.9490	0.9464
$(\alpha, \beta, \gamma) = (6, 0.5, 0.5)$									
n	20	30	60	90	100	120	250	350	450
$N.E.$	0.9500	0.9552	0.9468	0.9494	0.9534	0.9482	0.9492	0.9524	0.9574

Tab. 4.2. Niveau empirique pour $\alpha = (0.5, 1.5, 3, 6)$, β et $\gamma = 0.5$.

$(\alpha, \beta, \gamma) = (0.5, 1.5, 1.5)$									
n	20	30	60	90	100	120	250	350	450
$N.E.$	0.9586	0.9498	0.9484	0.9506	0.9536	0.9468	0.9532	0.9480	0.9506
$(\alpha, \beta, \gamma) = (1.5, 1.5, 1.5)$									
n	20	30	60	90	100	120	250	350	450
$N.E.$	0.9522	0.9574	0.9584	0.9478	0.9486	0.9494	0.9484	0.9552	0.9468
$(\alpha, \beta, \gamma) = (3, 1.5, 1.5)$									
n	20	30	60	90	100	120	250	350	450
$N.E.$	0.9492	0.9544	0.9618	0.9472	0.9540	0.9496	0.9482	0.9476	0.9520
$(\alpha, \beta, \gamma) = (6, 1.5, 1.5)$									
n	20	30	60	90	100	120	250	350	450
$N.E.$	0.9534	0.9467	0.9534	0.9498	0.9465	0.9522	0.9534	0.9468	0.9478

Tab. 4.3. Niveau empirique pour $\alpha = (0.5, 1.5, 3, 6)$, β et $\gamma = 1.5$.

$(\alpha, \beta, \gamma) = (0.5, 3, 3)$									
n	20	30	60	90	100	120	250	350	450
$N.E.$	0.9592	0.9518	0.9564	0.9506	0.9468	0.9444	0.9528	0.9492	0.9434
$(\alpha, \beta, \gamma) = (1.5, 3, 3)$									
n	20	30	60	90	100	120	250	350	450
$N.E.$	0.9516	0.9524	0.9576	0.9423	0.9498	0.9454	0.9494	0.9558	0.9478
$(\alpha, \beta, \gamma) = (3, 3, 3)$									
n	20	30	60	90	100	120	250	350	450
$N.E.$	0.9565	0.9512	0.9467	0.9487	0.9545	0.9534	0.9521	0.9540	0.9498
$(\alpha, \beta, \gamma) = (6, 3, 3)$									
n	20	30	60	90	100	120	250	350	450
$N.E.$	0.9528	0.9553	0.9590	0.9487	0.9477	0.9489	0.9487	0.9543	0.9469

Tab. 4.4. Niveau empirique pour $\alpha = (0.5, 1.5, 3, 6)$, β et $\gamma = 3$.

$(\alpha, \beta, \gamma) = (0.5, 8, 8)$									
n	20	30	60	90	100	120	250	350	450
$N.E.$	0.9419	0.9527	0.9520	0.9489	0.9482	0.9485	0.9460	0.9476	0.9551
$(\alpha, \beta, \gamma) = (1.5, 8, 8)$									
n	20	30	60	90	100	120	250	350	450
$N.E.$	0.9521	0.9556	0.9498	0.9456	0.9478	0.9490	0.9467	0.9572	0.9499
$(\alpha, \beta, \gamma) = (3, 8, 8)$									
n	20	30	60	90	100	120	250	350	450
$N.E.$	0.9487	0.9534	0.9507	0.9454	0.9460	0.9498	0.9478	0.9532	0.9455
$(\alpha, \beta, \gamma) = (6, 8, 8)$									
n	20	30	60	90	100	120	250	350	450
$N.E.$	0.9521	0.9543	0.9520	0.9488	0.9477	0.9489	0.9488	0.9523	0.9432

Tab. 4.5. Niveau empirique pour $\alpha = (0.5, 1.5, 3, 6)$, β et $\gamma = 8$.

Les tableaux 4.6 à 4.9 représentent le cas où le paramètre β varie.

$(\alpha, \beta, \gamma) = (0.25, 0.5, 0.25)$									
n	20	30	60	90	100	120	250	350	450
$N.E.$	0.9430	0.9450	0.9376	0.9442	0.9426	0.9417	0.9496	0.9558	0.9512
$(\alpha, \beta, \gamma) = (0.25, 1.5, 0.25)$									
n	20	30	60	90	100	120	250	350	450
$N.E.$	0.9462	0.9512	0.9511	0.9468	0.9456	0.9481	0.9493	0.9520	0.9512
$(\alpha, \beta, \gamma) = (0.25, 4, 0.25)$									
n	20	30	60	90	100	120	250	350	450
$N.E.$	0.9468	0.9549	0.9530	0.9508	0.9516	0.9458	0.9504	0.9490	0.9464
$(\alpha, \beta, \gamma) = (0.25, 8, 0.25)$									
n	20	30	60	90	100	120	250	350	450
$N.E.$	0.9500	0.9552	0.9468	0.9494	0.9534	0.9482	0.9492	0.9524	0.9574

Tab. 4.6. Niveau empirique pour $\beta = (0.25, 1.5, 4, 8)$, α et $\gamma = 0.25$.

$(\alpha, \beta, \gamma) = (2, 0.25, 2)$									
n	20	30	60	90	100	120	250	350	450
$N.E.$	0.9586	0.9487	0.9473	0.9517	0.9525	0.9476	0.9521	0.9480	0.9506
$(\alpha, \beta, \gamma) = (2, 1.5, 2)$									
n	20	30	60	90	100	120	250	350	450
$N.E.$	0.9533	0.9583	0.9573	0.9488	0.9489	0.9490	0.9494	0.9552	0.9468
$(\alpha, \beta, \gamma) = (2, 4, 2)$									
n	20	30	60	90	100	120	250	350	450
$N.E.$	0.9495	0.9564	0.9608	0.9463	0.9532	0.9493	0.9489	0.9476	0.9520
$(\alpha, \beta, \gamma) = (2, 8, 2)$									
n	20	30	60	90	100	120	250	350	450
$N.E.$	0.9524	0.9457	0.9524	0.9490	0.9475	0.9536	0.9520	0.9468	0.9478

Tab. 4.7. Niveau empirique pour $\beta = (0.25, 1.5, 4, 8)$, α et $\gamma = 2$.

$(\alpha, \beta, \gamma) = (5, 0.25, 5)$									
n	20	30	60	90	100	120	250	350	450
$N.E.$	0.9591	0.9523	0.9534	0.9516	0.9478	0.9474	0.9518	0.9493	0.9445
$(\alpha, \beta, \gamma) = (5, 1.5, 5)$									
n	20	30	60	90	100	120	250	350	450
$N.E.$	0.9526	0.9525	0.9559	0.9441	0.9495	0.9464	0.9492	0.9548	0.9488
$(\alpha, \beta, \gamma) = (5, 4, 5)$									
n	20	30	60	90	100	120	250	350	450
$N.E.$	0.9555	0.9532	0.9485	0.9493	0.9535	0.9534	0.9510	0.9543	0.9497
$(\alpha, \beta, \gamma) = (5, 8, 5)$									
n	20	30	60	90	100	120	250	350	450
$N.E.$	0.9518	0.9563	0.9593	0.9489	0.9487	0.9483	0.9482	0.9533	0.9479

Tab. 4.8. Niveau empirique pour $\beta = (0.25, 1.5, 4, 8)$, α et $\gamma = 5$.

$(\alpha, \beta, \gamma) = (10, 0.25, 10)$									
n	20	30	60	90	100	120	250	350	450
$N.E.$	0.9417	0.9537	0.9523	0.9484	0.9482	0.9484	0.9461	0.9477	0.9541
$(\alpha, \beta, \gamma) = (10, 1.5, 10)$									
n	20	30	60	90	100	120	250	350	450
$N.E.$	0.9525	0.9557	0.9493	0.9466	0.9468	0.9493	0.9457	0.9570	0.9499
$(\alpha, \beta, \gamma) = (10, 4, 10)$									
n	20	30	60	90	100	120	250	350	450
$N.E.$	0.9467	0.9434	0.9517	0.9454	0.9469	0.9478	0.9488	0.9522	0.9475
$(\alpha, \beta, \gamma) = (10, 8, 10)$									
n	20	30	60	90	100	120	250	350	450
$N.E.$	0.9526	0.9547	0.9544	0.9480	0.9467	0.9474	0.9485	0.9513	0.9472

Tab. 4.9. Niveau empirique pour $\beta = (0.25, 1.5, 4, 8)$, α et $\gamma = 10$.

Les tableaux 4.10 à 4.13 représentent le cas où le paramètre γ varie.

$(\alpha, \beta, \gamma) = (0.75, 0.75, 1.25)$									
n	20	30	60	90	100	120	250	350	450
$N.E.$	0.9522	0.9542	0.9487	0.9534	0.9518	0.9526	0.9448	0.9568	0.9502
$(\alpha, \beta, \gamma) = (0.75, 0.75, 2)$									
n	20	30	60	90	100	120	250	350	450
$N.E.$	0.9542	0.9544	0.9527	0.9449	0.9444	0.9475	0.9479	0.9525	0.9501
$(\alpha, \beta, \gamma) = (0.75, 0.75, 5)$									
n	20	30	60	90	100	120	250	350	450
$N.E.$	0.9556	0.9566	0.9546	0.9560	0.9542	0.9508	0.9552	0.9491	0.9494
$(\alpha, \beta, \gamma) = (0.75, 0.75, 9)$									
n	20	30	60	90	100	120	250	350	450
$N.E.$	0.9510	0.9551	0.9418	0.9491	0.9544	0.9482	0.9496	0.9554	0.9544

Tab. 4.10. Niveau empirique pour $\gamma = (1.25, 2, 5, 9)$, α et $\beta = 0.75$.

$(\alpha, \beta, \gamma) = (1.75, 1.75, 1.25)$									
n	20	30	60	90	100	120	250	350	450
$N.E.$	0.9586	0.9498	0.9484	0.9506	0.9536	0.9468	0.9532	0.9480	0.9506
$(\alpha, \beta, \gamma) = (1.75, 1.75, 2)$									
n	20	30	60	90	100	120	250	350	450
$N.E.$	0.9522	0.9574	0.9584	0.9478	0.9486	0.9494	0.9484	0.9552	0.9468
$(\alpha, \beta, \gamma) = (1.75, 1.75, 5)$									
n	20	30	60	90	100	120	250	350	450
$N.E.$	0.9492	0.9544	0.9618	0.9472	0.9540	0.9496	0.9482	0.9476	0.9520
$(\alpha, \beta, \gamma) = (1.75, 1.75, 9)$									
n	20	30	60	90	100	120	250	350	450
$N.E.$	0.9534	0.9467	0.9534	0.9498	0.9465	0.9522	0.9534	0.9468	0.9478

Tab. 4.11. Niveau empirique pour $\gamma = (1.25, 2, 5, 9)$, α et $\beta = 1.75$.

$(\alpha, \beta, \gamma) = (3.5, 3.5, 1.25)$									
n	20	30	60	90	100	120	250	350	450
$N.E.$	0.9492	0.9518	0.9574	0.9516	0.9468	0.9544	0.9528	0.9498	0.9499
$(\alpha, \beta, \gamma) = (3.5, 3.5, 2)$									
n	20	30	60	90	100	120	250	350	450
$N.E.$	0.9546	0.9544	0.9576	0.9423	0.9495	0.9455	0.9495	0.9558	0.9478
$(\alpha, \beta, \gamma) = (3.5, 3.5, 5)$									
n	20	30	60	90	100	120	250	350	450
$N.E.$	0.9555	0.9512	0.9447	0.9485	0.9555	0.9535	0.9521	0.9543	0.9498
$(\alpha, \beta, \gamma) = (3.5, 3.5, 9)$									
n	20	30	60	90	100	120	250	350	450
$N.E.$	0.9538	0.9554	0.9593	0.9437	0.9473	0.9488	0.9489	0.9523	0.9479

Tab. 4.12. Niveau empirique pour $\gamma = (1.25, 2, 5, 9)$, α et $\beta = 3.5$.

$(\alpha, \beta, \gamma) = (6, 6, 1.25)$									
n	20	30	60	90	100	120	250	350	450
$N.E.$	0.9419	0.9527	0.9520	0.9489	0.9482	0.9485	0.9460	0.9476	0.9551
$(\alpha, \beta, \gamma) = (6, 6, 2)$									
n	20	30	60	90	100	120	250	350	450
$N.E.$	0.9521	0.9556	0.9498	0.9456	0.9478	0.9490	0.9467	0.9572	0.9499
$(\alpha, \beta, \gamma) = (6, 6, 5)$									
n	20	30	60	90	100	120	250	350	450
$N.E.$	0.9487	0.9534	0.9507	0.9454	0.9460	0.9498	0.9478	0.9532	0.9455
$(\alpha, \beta, \gamma) = (6, 6, 9)$									
n	20	30	60	90	100	120	250	350	450
$N.E.$	0.9521	0.9543	0.9520	0.9488	0.9477	0.9489	0.9488	0.9523	0.9492

Tab. 4.13. Niveau empirique pour $\gamma = (1.25, 2, 5, 9)$, α et $\beta = 6$.

Une lecture des tableaux 4.2 à 4.13, nous indique que les valeurs du niveau empirique N.E calculées sont très proches de la valeur du niveau théorique N.E=0.95. Nous pouvons conclure que le test proposé est bien adapté à la distribution IWG.

4.2 Simulation du modèle AFT-IWG

4.2.1 Estimateurs du maximum de vraisemblance des paramètres du modèle AFT-IWG

Pour confirmer la propriété de la convergence des estimateurs du maximum de vraisemblance, nous calculons l'erreur quadratique moyenne pour les EMV simulés des paramètres du modèle AFT-IGW.

Supposons que le modèle AFT-IGW est considéré. Les données ont été simulées $N = 10.000$ fois, avec les valeurs suivantes des paramètres : ($\alpha = 2, \nu = 3, \gamma = 2, \beta_0 = 6, \beta_1 = 0,82$). Nous utilisons le logiciel statistique R et l'algorithme Barzilai-Borwein (BB) (Ravi (2009)). Les valeurs de la moyenne des estimateurs du maximum de vraisemblance $\hat{\alpha}, \hat{\nu}, \hat{\gamma}, \hat{\beta}_0$ et $\hat{\beta}_1$ et leurs **erreurs quadratiques moyennes (SME)** sont calculées et données dans le tableau 4.14 (taille des échantillons $n = 150, n = 350$, et $n = 550$).

$N = 10,000$	$n = 150$	$n = 350$	$n = 550$
$\hat{\alpha}$	2.1273	2.1115	2.0933
SME	0.0021	0.0015	0.0006
$\hat{\nu}$	2.9577	3.0847	3.0251
SME	0.0035	0.0021	0.0007
$\hat{\gamma}$	2.1125	2.0887	2.0252
SME	0.0045	0.0026	0.0004
$\hat{\beta}_0$	6.0278	6.0689	5.9882
SME	0.0053	0.0074	0.0006
$\hat{\beta}_1$	0.8710	0.8425	0.8373
SME	0.0045	0.0038	0.0004

Tab. 4.14. Erreurs quadratiques moyennes des estimateurs du maximum de vraisemblance $\hat{\alpha}, \hat{\nu}, \hat{\gamma}, \hat{\beta}_0$ et $\hat{\beta}_1$.

Les résultats de la simulation confirment le fait bien connu que pour les modèles réguliers, **les estimateurs du maximum de vraisemblance sont convergents**.

La figure 4.3 montre que tous les estimateurs du maximum de vraisemblance des paramètres du modèle AFT-IWG convergent plus vite que $n^{-0,5}$, ce qui confirme que **les estimateurs du maximum de vraisemblance du modèle AFT-IWG sont \sqrt{n} -convergents**.

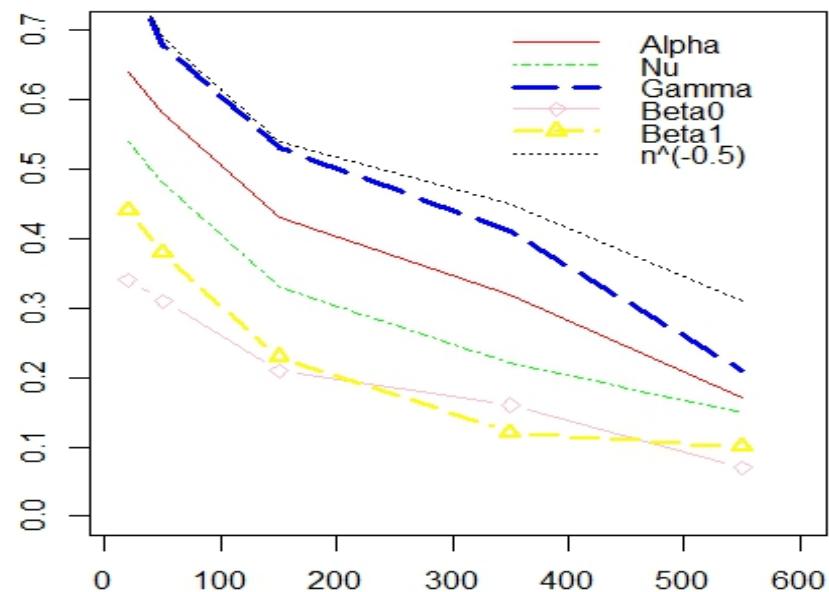


Fig. 4.3 Erreurs absolues moyennes des EMV $\hat{\alpha}$, $\hat{\nu}$, $\hat{\gamma}$, $\hat{\beta}_0$ et $\hat{\beta}_1$, simulés $N = 10,000$ fois.

4.2.2 Distribution de la statistique Y_n^2

Nous simulons $N = 10.000$ fois des échantillons du modèle AFT-IWG, avec différentes valeurs des paramètres et avec des nombres différents de groupement de classes k .

Nous présentons dans la figure 4.4, les histogrammes des échantillons simulés par rapport à la distribution du Chi-deux avec $k = 10$ degrés de liberté.

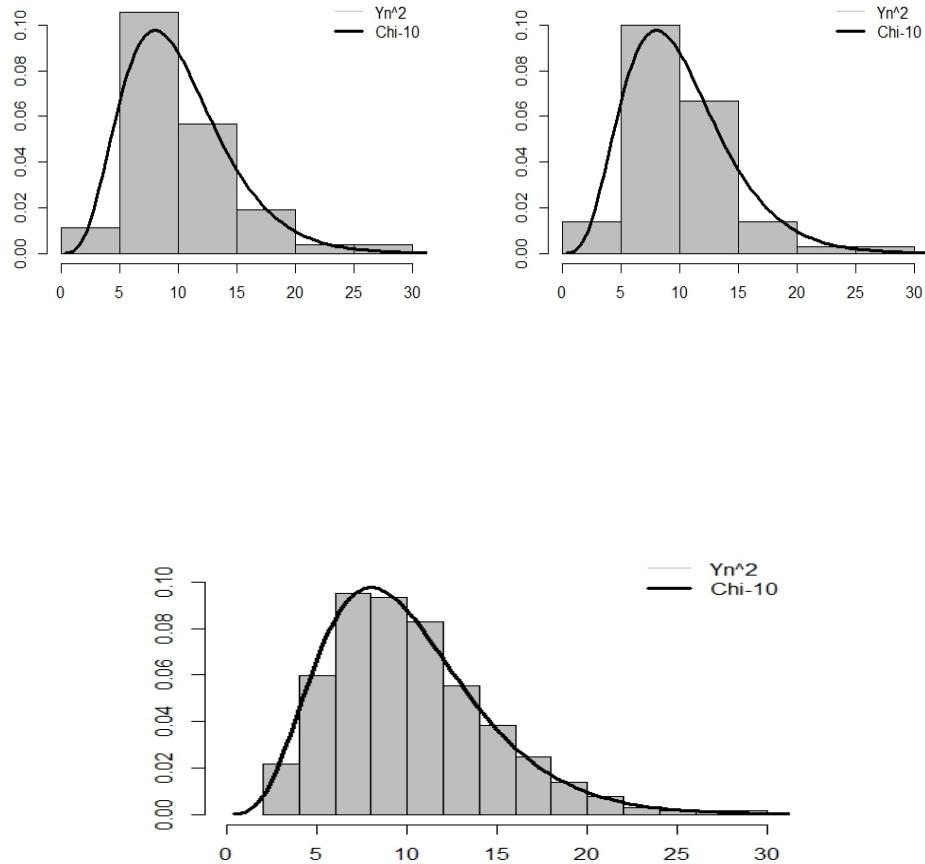


Fig. 4.4 Histogrammes de Y_n^2 pour $(\hat{\alpha} = 2, \hat{\nu} = 3, \hat{\gamma} = 2, \hat{\beta}_0 = 6, \hat{\beta}_1 = 0.82)$,
et $(k = 10, n = 50, 150, 350.)$

Comme prévu, la statistique Y_n^2 du test NRR pour le modèle AFT-IWG suit à la limite, la distribution du Chi-deux avec k degrés de liberté. On conclut que la statistique Y_n^2 du test NRR est invariante.

4.2.3 La statistique du test Y_n^2 pour le modèle AFT-IWG

Nous simulons $N = 10,000$ échantillons de tailles respectives $n = 50, n = 150, n = 350$ et $n = 550$, suivant le modèle AFT-IWG. Puis, nous calculons la statistique Y_n^2 pour chaque échantillon. Ensuite, on compte le nombre de cas de rejet de l'hypothèse nulle H_0 .

Le tableau 4.15, représente pour chaque risque d'erreur α ($\alpha = 0.01, \alpha = 0.05$ et $\alpha = 0.1$), les résultats des niveaux de signification de la statistique Y_n^2 du test NRR, contre leurs valeurs théoriques.

$N = 10,000$	$\alpha = 0.01$	$\alpha = 0.05$	$\alpha = 0.1$
$n = 50$	0.0089	0.0486	0.0924
$n = 150$	0.0096	0.0489	0.0957
$n = 350$	0.0124	0.0512	0.1305
$n = 550$	0.0112	0.0504	0.1102

Tab. 4.15. Niveaux de significations simulés de $Y_n^2(\hat{\theta})$, contre leurs valeurs théoriques.

De ces résultats, et en tenant compte des erreurs de simulation, nous observons que les niveaux de signification simulés de la statistique Y_n^2 du test NRR, coïncident avec ceux correspondant aux niveaux théoriques de la distribution du Chi-deux avec k (ici $k = 10$) degrés de liberté.

Par conséquent, nous pouvons dire que la statistique Y_n^2 du test NRR proposée peut ajuster des données du modèle AFT inverse Weibull généralisé (AFT-IWG) de manière très satisfaisante.

4.3 Applications à la distribution IWG

Dans cette section nous appliquons les tests construits dans ce travail à des données réelles de fiabilité et d'analyse de survie pour montrer l'efficacité et la maniabilité des tests proposés et leur applicabilité dans différents domaines d'étude.

4.3.1 Exemple de données de survie

Les données de "Arm A" du cancer de la tête et du cou [Efron (1988)], représentant les durées de survie en jours pour $n = 51$ patients, sont données dans le tableau 4.16, où * représente la censure.

7, 34, 42, 63, 64, 74*, 83, 84, 91, 108, 112, 129, 133, 133, 139, 140, 140, 146, 149, 154, 157, 160, 160, 165, 173, 176, 185*, 218, 225, 241, 248, 273, 277, 279*, 297, 319*, 405, 417, 420, 440, 523*, 523, 583, 594, 1101, 1116*, 1146, 1226*, 1349*, 1412*, 1417.
--

Tab. 4.16. Données d'Efron.

En transformant ces données en mois, Efron (1988) a estimé la fonction de hasard pour ces données en montrant qu'elle est unimodale. Mudholkar et al. (1996) ont montré que la distribution de Weibull Exponentielle peut ajuster ces données. Nikulin et Haghighi (2004, 2009) ont montré, en utilisant le test d'Akritis, que la distribution de Weibull généralisée donne un bon ajustement pour ces données. Saïdia et Seddik ameur (2010) ont montré que la loi Gaussienne inverse ajuste beaucoup mieux ces données. Nous nous sommes proposées de l'ajuster par une distribution IWG.

Pour la distribution inverse Weibull généralisée (IWG), l'estimateur du vecteur de paramètres $\hat{\theta}$ de θ maximisant la fonction de log-vraisemblance est :

$$\hat{\theta} = (\hat{\alpha}, \hat{\beta}, \hat{\gamma})^T = (19.1894, 4.6986, 3.7825)^T, \quad \text{log-vrais} = -2.9757.$$

Nous choisissons un nombre de classes $k = 5$, les éléments de la statistique Y_n^2 du test NRR sont présentés comme suit :

\widehat{a}_j	3.7608	6.0502	11.5407	23.0934	47, 24
U_j	9	15	5	6	16
e_j	9.1206	19.1206	9.1206	9.1206	9.1206
\widehat{C}_{1j}	-0.00523	0.0049	0.0108	0.0130	0.1340
\widehat{C}_{2j}	0.0501	0.0813	0.0241	0.0513	0.1036
\widehat{C}_{3j}	0.0154	-0.0610	-0.1125	-0.0711	-0.0647

La matrice d'information de Fisher est :

$$\widehat{I}_{3 \times 3} = \begin{pmatrix} -0.7999923 & -2.781990 & 0.6746561 \\ -2.781990 & 7.316385 & -0.7999923 \\ 0.6746561 & -0.7999923 & 1.5894896 \end{pmatrix}$$

Après calcul, nous trouvons $Y_n^2 = 10.8094$.

Et la statistique de test de Habib et Thomas est $Y_n^2(\widehat{\theta}) = 11.0694$.

Au seuil $\alpha = 0.05$, la valeur critique $\chi_5^2 = 11.07050$, d'où la distribution inverse Weibull généralisée IWG donne aussi un bon ajustement pour les données d'Efron "Arm A".

4.3.2 Exemple de fiabilité

Ces données tirées du livre de Bagdonavičius et al. (2010), représentent les données de $n = 120$ dispositifs électroniques qui ont été observées pour $\tau = 5.54$ (années).

Le nombre de défaillances est $\delta = 113$.

5.5199	5.5335	5.5232	5.4860	5.4842	3.8848	3.7823	3.4684
5.1816	5.2546	5.4360	5.0929	3.4382	3.5992	3.7502	3.5019
5.3353	5.2441	5.4291	4.6324	3.5846	3.6574	3.4438	2.6205
5.4633	5.0898	5.4903	4.8858	3.5587	3.6674	3.7062	2.8370
5.0144	5.0517	5.1219	4.6001	3.3514	3.654	3.1586	2.8373
5.0951	5.1223	4.8619	3.5363	3.3625	3.7288	3.1730	2.8766
4.8406	4.8635	5.1023	3.7157	3.3802	3.3492	3.2132	2.9888
4.8679	4.9466	4.8928	3.4012	3.3855	3.3507	3.2323	3.0720
4.5992	4.6400	4.8532	3.5297	3.5110	3.8902	3.9113	3.9468
4.8351	4.8009	4.7300	3.9787	3.9903	4.0078	4.0646	4.1301
5.1766	4.9846	4.7881	4.5826	4.2919	4.2970	4.2312	4.2300
5.1710	5.0008	4.7969	4.5388	4.4932	4.2581	4.3666	4.4365
2.1461	2.3858	2.5404	3.9728	4.2885	1.9172	4.2525	4.1427
3.9551	2.3079	1.7440	4.3918	4.4919	2.3753	2.4147	2.6471
4.7164							

Supposons que les taux de défaillance suivent une distribution inverse Weibull généralisée (IWG), les paramètres du modèle sont estimés par :

$$\hat{\alpha} = 0.0241, \hat{\beta} = 1.1034 \text{ et } \hat{\gamma} = 2.3107.$$

Nous choisissons de prendre 10 intervalles, c'est à dire $k = 10$. Les résultats intermédiaires pour calculer la statistique Y_n^2 sont présentés ci-dessous :

\hat{a}_j	3.8402	4.0125	4.5482	5.1182	5.5335
U_j	29	8	17	26	17
e_j	2.3046	2.3046	2.3046	2.3046	2.3046
\widehat{C}_{1j}	-0.04833	0.0689	0.2508	0.2430	0.1400
\widehat{C}_{2j}	0.1521	0.0873	0.0873	0.0504	0.0816
\widehat{C}_{3j}	0.0124	-0.0700	-0.1046	-0.0851	-0.0787

La matrice d'information de Fisher estimée est obtenue comme suit :

$$\widehat{I}_{3 \times 3} = \begin{pmatrix} -1.4453713 & 0.6635231 & -1.3748646 \\ 0.6635231 & 0.6868031 & -0.4618151 \\ -1.3748646 & -0.4618151 & 1.9582996 \end{pmatrix}$$

La matrice \hat{W} :

$$\hat{W} = (\hat{W}_1, \hat{W}_2, \hat{W}_3)^T = (0.5860, -0.0942, 1.2564)^T$$

Après calcul, nous remarquons que dans le cas de la distribution IWG la matrice \hat{G} est dégénérée. Donc le rang de la matrice Σ est $k - 1 = 9$.

La valeur de la statistique de test est :

$$Y_n^2 = X^2 + Q = 13,2145 + 6,7572 = 19,9997.$$

Et la valeur de la statistique de Habib et Thomas est :

$$Y_n^2(\hat{\theta}) = 21.874$$

Donc, à partir de ces résultats, nous pouvons dire que les taux de défaillance des dispositifs électroniques ne suivent pas la distribution inverse Weibull généralisée (IWG).

4.4 Applications au modèle AFT-IWG

Cette fois, nous considérons des exemples qui représentent des données d'analyse de survie et de fiabilité auxquelles nous allons appliquer le test construit précédemment pour vérifier l'hypothèse que ces données suivent un modèle AFT-IWG .

4.4.1 Exemples de données de survie

Exemple de données de lupus

Nous étudions l'exemple des données de survie censurées, extraites du fichier "lupus" sur le site web : <http://lib.stat.cmu.edu/datasets/>. Les données représentent les durées de survie de 40 patients de la néphrite lupus, qui sont les temps de survie (en mois) de 40 patients atteints de lupus ayant subi une biopsie rénale entre 1967 et 1983 et qui ont été suivis jusqu'à la mort ou à la fin de 1990.

Nous appliquons le test NRR à ces données, pour vérifier si elle peuvent être ajustées par un modèle AFT inverse Weibull généralisé.

x_i	157*	268*	209	21	28	46	169	39	99	132	114	127	103
z_{i1}	1	10	2	0.1	7	24	2	22	0.1	0.1	1	4	11
x_i	39	63	159	75	32	148*	138	145	182	4	103	180	54
z_{i1}	57	1	0.1	1	1	44	10	0.1	14	106	1	7	1
x_i	188	19	118	48	13	165	86	78	25	4	65	62	89*
z_{i1}	0.1	0.1	10	1	6	0.1	8	0.1	15	34	0.1	8	0.1
													77

* données censurées.

En utilisant le logiciel statistique R, nous trouvons les valeurs des estimateurs du maximum de vraisemblance des paramètres :

$$\hat{\alpha} = 0.56002, \hat{\nu} = 0.0042, \hat{\gamma} = 0.0797, \hat{\beta}_0 = -17.5789, \hat{\beta}_1 = 37.9916.$$

Nous prenons (par exemple) 6 intervalles ($k = 6$).

Pour calculer la statistique Y_n^2 du test NRR, nous présentons les calculs intermédiaires :

\hat{a}_j	28.46	60.2	94.2	136.78	172.56	268
U_j	7	7	7	7	7	5
Z_j	-1.584	-1.584	-1.584	-1.584	-1.584	-1.9
e_j	17.019	17.019	17.019	17.019	17.019	17.019

La matrice estimée de la matrice C est :

$$\widehat{C} = \begin{pmatrix} -1.272 * 10^{-3} & 1.186 * 10^{-3} & 5.337 * 10^{-5} & 1.231 * 10^{-3} & 1.257 * 10^{-3} & -0.006 \\ -42.319 & 40.608 & 1.130 & 41.397 & 41.690 & -208.690 \\ 0.045 & -0.119 & 0.046 & -0.081 & -0.059 & 0.276 \\ 1.535 * 10^{-5} & -4.017 * 10^{-5} & 1.549 * 10^{-5} & -2.722 * 10^{-5} & -1.993 * 10^{-5} & 9.346 * 10^{-5} \\ -1.065 * 10^{-4} & -9.631 * 10^{-5} & 1.365 * 10^{-4} & 00 & 3.68 * 10^{-5} & -2.15 * 10^{-4} \end{pmatrix}$$

D'où

$$\widehat{W} = [0.0736 \quad 2425.971 \quad -2.6921 \quad -0.0009 \quad 0.0035]^T,$$

et la matrice d'information de Fisher estimée

$$\widehat{I}_{5 \times 5} = \begin{pmatrix} 4.28 * 10^{-4} & 14.501 & -3.669 * 10^{-2} & -1.183 * 10^{-5} & -2.2282 * 10^{-5} \\ 14.501 & 4.911 * 10^{+5} & -1219 & -0.412 & -0.756 \\ -3.669 * 10^{-2} & -1219 & 3.070 & 1.037 * 10^{-3} & 1.982 * 10^{-3} \\ -1.183 * 10^{-5} & -0.412 & 1.037 * 10^{-3} & 00 & 00 \\ -2.2282 * 10^{-5} & -0.756 & 1.982 * 10^{-3} & 00 & 00 \end{pmatrix}$$

Donc on trouve

$$\widehat{G} = \begin{pmatrix} 7.496 * 10^{-5} & 2.797 & -1.99 * 10^{-2} & 00 & -3.356 * 10^{-5} \\ 2.797 & 1.033 * 10^{+5} & -686 & -0.2318 & -1.129 \\ -1.991 * 10^{-2} & -686 & -2.296 & 7.759 * 10^{-4} & 2.396 * 10^{-3} \\ 00 & -0.2318 & 7.759 * 10^{-4} & 00 & 00 \\ -3.356 * 10^{-5} & -1.129 & 2.396 * 10^{-3} & 00 & 00 \end{pmatrix}$$

Nous obtenons les quantités

$$Q = 3.306, \quad X^2 = 8.986$$

la valeur de la statistique de test est alors $Y_n^2 = 12.292 < \chi_{0.05,6}^2 = 12.5916$.

D'où, nous concluons que la distribution AFT-GIW peut s'adapter au données de "lupus".

La statistique du test Habib et Thomas pour ces données est :

$$Y_n^2(\hat{\theta}) = 11.025$$

Ce qui signifie que le modèle AFT-IWG ajuste les données de Lupus selon les deux tests.

Exemple de Lawless (Cancer du poumon)

Nous utilisons les données de Lawless (2003) qui représentent les données de 40 patients de cancer du poumon. Chaque patient est indiqué par un statut de performance (PS) nommé l'indice de Karnofsky, sur l'échelle de 10 à 90 : 10, 20 et 30, 60 s'il est limité à l'hôpital, et 70, 80 et 90 s'il est en mesure de prendre soin de lui-même. Donc, nous utilisons deux variables : le Traitement et l'état de la performance.

L'objectif principal de l'étude initiale était de comparer les effets de deux traitements de chimiothérapie à prolonger le temps de survie. Tous les patients ont reçu un traitement préalable et ont ensuite été assignés au hasard à l'un des deux traitements, standard et test, présentés dans le tableau suivant :

Traitement : standard, * censuré

Temps	411	126	118	82	8	25*	11
PS	70	60	70	40	40	70	70
Temps	54	153	16	56	21	287	10
PS	80	60	30	80	40	60	40
Temps	8	12	177	12	200	250	100
PS	20	50	50	40	80	70	60

Traitements : test, * censuré

Temps	999	231*	991	1	201	44	15
PS	90	50	70	20	80	60	50
Temps	103*	2	20	51	18	90	84
PS	70	40	30	30	40	60	80
Temps	164	19	43	340	231		
PS	70	30	60	80	70		

Si on choisit $k = 6$ intervalles, les estimateurs du maximum de vraisemblance des paramètres du modèle AFT-IWG obtenus, sont :

$$\hat{\alpha} = 3.5641, \hat{\nu} = 4.0172, \hat{\gamma} = 2.1348,$$

$$\hat{\beta}_0 = 0.9843, \hat{\beta}_1 = 0.0686, \hat{\beta}_2 = 0.2863.$$

Pour appliquer la statistique Y_n^2 du test NRR à ces données, nous avons calculé les éléments suivants :

\hat{a}_j	10.54	30.94	65.54	119.66	218.01	999
U_j	5	9	5	5	6	7
\hat{e}_j	6.3468	6.3468	6.3468	6.3468	6.3468	6.3468

La matrice inverse de la matrice \hat{G} est

$$\hat{G}^{-1} = \begin{pmatrix} 0.0021 & 0.0064 & 0.021 & 0.0461 & 0.0845 & 0.0905 \\ 0.0064 & 0.0054 & 0.1346 & 0.3408 & 0.4107 & 0.6751 \\ 0.021 & 0.1346 & 0.0087 & 0.6432 & 0.7655 & 0.9921 \\ 0.0461 & 0.3408 & 0.6432 & 0.0076 & 0.0213 & 0.9321 \\ 0.0845 & 0.4107 & 0.7655 & 0.0213 & 0.0092 & 0.8675 \\ 0.0905 & 0.6751 & 0.9921 & 0.9321 & 0.8675 & 0.0089 \end{pmatrix}$$

Le vecteur \hat{W} est donné par :

$$\hat{W} = (2.1304, -0.0021, -0.0543, 1.5634, 1.8764, -0.00436)^T,$$

Après calcul, on trouve :

$$Q = 0.5638,$$

et

$$X^2 = 2.8643,$$

donc la statistique $Y_n^2 = 3.4281$.

D'un autre côté, la valeur de la statistique du test de Habib et Thomas est : $Y_n^2(\hat{\theta}) = 5.817$.

Nous concluons que L'hypothèse testée selon laquelle les données du cancer du poumon sont ajustées par le modèle AFT-IWG n'est pas rejetée.

4.4.2 Exemple de données de fiabilité

Nous considérons l'exemple utilisé par Bagdonavičius et Nikulin (2011), donnant le temps de rupture électrique d'unités testées à 100.3, 122.4, 157.1, 219.0 et 361, 4KV/mm, pour évaluer la structure du système des tensions.

Les principaux objectifs de l'analyse étaient de lier le temps de panne et le voltage et d'obtenir un modèle qui pourrait être utilisé pour l'interpolation à des voltages inférieurs. Les méthodes de diagnostic suggèrent un modèle AFT-Weibull. Lawless a utilisé une des méthodes paramétriques basées sur les rapports de vraisemblance pour vérifier la distribution Weibull et le modèle AFT-Weibull power rule séparément. Aucun des tests ne contredit le modèle AFT-Weibull power rule. Mais Bagdonavičius et Nikulin (2011) ont montré (en utilisant le test d'ajustement de Chi-deux modifié) que ces données ne suivent pas le modèle AFT-Weibull power rule.

Nous appliquons maintenant à ces données un test d'ajustement du Chi-deux au modèle AFT-IWG.

Temps de rupture diélectrique d'unités testées	Le stress de Voltage kV/mm
606, 1012, 2520, 2610, 3988, 4100, 5025, 6842	100.3
188, 297, 405, 744, 1218, 1340, 1715, 3382	122.4
49, 99, 154.5, 180, 291, 447, 510, 600, 1656, 1721	157.1
15, 16, 36, 50, 55, 95, 122, 129, 625, 700	219.0
0.10, 0.33, 0.50, 0.90, 1.00, 1.55, 1.65, 2.10, 4.00	361.4

En utilisant le logiciel statistique R, nous trouvons les valeurs des estimateurs du maximum de vraisemblance des paramètres :

$$\hat{\alpha} = 1.4767, \hat{\nu} = 2.4607 * 10^{-7}, \hat{\gamma} = 1.3831, \\ \hat{\beta}_0 = -9.2415, \hat{\beta}_1 = 27.3021.$$

Si on choisit par exemple $k = 6$ intervalles, pour calculer la statistique Y_n^2 du test NRR, nous avons besoin de ces résultats :

\hat{a}_j	3.85	90.08	295.4	687.43	2328.56	6842
U_J	8	7	8	7	8	7
Z_j	0.869	0.72	0.869	0.72	0.869	0.72
e_j	2.166	2.166	2.166	2.166	2.166	2.166
\hat{C}_{4j}	$-3.7 * 10^{-8}$	$-3.24 * 10^{-8}$	$-3.7 * 10^{-8}$	$-3.24 * 10^{-8}$	$-3.7 * 10^{-8}$	$-3.24 * 10^{-8}$
\hat{C}_{5j}	$-3.69 * 10^{-10}$	$-2.65 * 10^{-10}$	$-2.44 * 10^{-10}$	$-1.73 * 10^{-10}$	$-1.52 * 10^{-10}$	$-8.96 * 10^{-11}$

Le vecteur \hat{W} est alors :

$$\hat{W} = (-1.2034, 0.0021, -0.0534, 1.0374, -1.7742, 0.05276)^T,$$

et la forme quadratique Q de la statistique Y_n^2 :

$$Q = 45.8274,$$

et

$$X^2 = 22.7813$$

On déduit la statistique Y_n^2 du test NRR est : $Y_n^2 = 68,6087$.

Il est évident que ces données contredisent, le fait quelles suivent un modèle AFT-IWG.

Mais, en appliquant le test de Habib et Thomas, la valeur de la statistique : $Y_n^2(\hat{\theta}) = 9,712$. Donc le test ne contredit pas l'hypothèse H_0 .

Conclusion et perspectives

Dans ce travail, nous avons construit un modèle de durées de vie très flexible, soumis à un temps de vie accéléré, par des variables explicatives représentant différents stress. Ce type de modèle est très utile particulièrement en industrie, car il permet d'optimiser la fiabilité du matériel et sa maintenance.

En analyse de survie, il peut aussi donner les résultats de certains essais thérapeutiques en un temps relativement court.

Nous avons par ailleurs, construit un test d'ajustement pour ce modèle, en tenant compte de la censure. Nous avons donné la forme explicite de la statistique de test et nous avons élaboré un logiciel de calcul nous permettant de vérifier la validité du modèle. Nous avons aussi calculé complètement la matrice d'information de Fisher du modèle.

En perspectives, nous nous proposons de faire une étude comparative de ce modèle et des modèles Weibull généralisé, inverse gaussien et logistique, et ceci selon la forme de la fonction risque. Nous projetons aussi, l'étude de mélange de deux lois inverses Weibull généralisées, pour décrire une même population avec des paramètres différents.

liste de communications

- Goual Hafida & Seddik-Ameur Nacira, Chi-Squared Type Test for the AFT-Generalized Inverse Weibull Distribution, Communications in Statistics- Theory and Methods, 43, 13, 2605-2617.

-Goual Hafida & Seddik-Ameur Nacira, Modeling un AFT-generalized inverse Weibull model in case of censored data, Congrès des mathématiciens algériens, Mai 11-13, 2014, Tlemcen, Algérie.

-Goual Hafida & Seddik-Ameur Nacira, An application of the AFT model based on the generalised inverse Weibull distribution for reliability analysis, 19ème colloque CSMT2014, Mars 17-20, 2014, Sousse, Tunisie.

-Goual Hafida & Seddik-Ameur Nacira, La modélisation de la distribution de l'inverse Weibull généralisée et sa propre application, Conférence sur les problèmes mathématiques non linéaires ProMa'2014, 11-13 Mars, 2014, Annaba, Algérie.

-Goual Hafida & Seddik-Ameur Nacira, Comparison of proportional hazard models and models of accelerated failure times ; Application to real data of lung cancer, 5th international conference on modeling, simulation and applied optimization, ICMSAO 13, April 28-30, 2013, Hammamet, Tunisia.

-Goual Hafida & Seddik-Ameur Nacira, On the Statistics of the Generalized Inverse Weibull Distribution, Journées Jeunes Chercheurs JJC'13, Avril 22-23, 2013, Annaba, Algérie.

-Goual Hafida & Seddik-Ameur Nacira, The Generalised Inverse Wibull Distribution, le premier colloque international sur les équations différentielles, CIED2013, Mars 5-8, 2013, Ghardaia, Algerie.

-Goual Hafida & Seddik-Ameur Nacira, The accelerated failure time model against Cox model, Journées d'Analyse Mathématique et Applications, Mars 22-31, 2012, Hammamet, Tunisie.

Bibliographie

- [1] Abernathy, R. B. (2004) The New Weibull Handbook, 4th Edition, Dept. AT Houston, Texas, USA 77252-2608.
- [2] Aguirre, N., Nikulin, M. (1994) Chi-Squared Goodness-of-fit Test for the Family of Logistic Distribution, *Kybernetika*, 30 : 214-222.
- [3] Akritas, Michael G. (1988) Pearson-Type Goodness-of-Fit Tests : The Univariate Case. *Journal of the American Statistical Association*, 83 (401) :222-230.
- [4] Aleem, M and Pasha G.R (2003). Ratio, product and single Moments of Order Statistics from Inverse Weibull Distribution. *J.Stat. Vol.* 10, (1) 1-7.
- [5] Andersen PK. (1982) Testing goodness of fit if cox's regression and life model. *Biometrics* 38, p 67-77.
- [6] Andersen PK. (1991) Survival analysis 1982-1991 : the second decade of the proportional hazards regression model. *Statistics in medicine* 10, p 1931-1941.
- [7] Bagdonavičius, V., (1978). Testing the hypothesis of the additive accumulation of dam-ages. *Probab. Theory and its Appl.*, 23 (2) , 403-408.
- [8] Bagdonavičius, V., Cheminade, O., Nikulin, M. (2004) Statistical planning and inference in accelerated life testing using the CHSS model, *Journal of the Statistical Planning and Inference*, 2 : 535-551.
- [9] Bagdonavičius, V., Clerjaud, L., Nikulin, M. (2008) Accelerated Life Testing when the Hazard Rate Function has Cup Shape, In : Mathematical methods in Survival analysis, reliability and quality of life (Huber

C., Limnios N., Balakrishnan N., Messbah M., Nikulin M. (Eds)), ISTE & Wiley : London, 203-215.

- [10] Bagdonavičius, V., Kruopis, J., Nikulin, M. (2011) Non-parametric Tests for Censored Data, John Wiley & Sons Inc.
- [11] Bagdonavičius, V., Masiulaityle, I., Nikulin, M. (2008a) Statistical analysis of redundant system "warm" stand-by units, Stochastics : An International Journal of Probability and Stochastic Processes, 80 #2-3, 115-128.
- [12] Bagdonavičius, V., Masiulaityle, I., Nikulin, M. (2008b) Statistical analysis of redundant system with one stand-by unit, In : Mathematical methods in Survival analysis, reliability and quality of life (Huber C., Limnios N., Balakrishnan N., Messbah M., Nikulin M. (Eds)), ISTE & Wiley : London, 189-202 .
- [13] Bagdonavičius, V., Masiulaityle, I., Nikulin, M. (2009) Advances in Degradation Modeling : Applications to Reliability, Survival Analysis, and Finance (Nikulin,M., Limnios N., Balakrishnan N., Kahle W., Huber C. (Eds)), Birkhäuser Boston, 293-310.
- [14] Bagdonavičius, V., Masiulaityle, I., Nikulin, M. (2010) Parametric estimation of redundant system reliability from censored data. In : Mathematical and Statistical Methods in Reliability (Balakrishnan, N., Nikulin,M., Rykov, V. (Eds)), Springer Verlag : Boston.
- [15] Bagdonavičius V., Nikulin, M. (2002) Accelereted Life Models : Modeling and Statistical Analysis, Chapman and Hall.
- [16] Bagdonavičius, V., Nikulin, M. (2011) Chi-squared tests for general composite hypotheses from censored samples Comptes Rendus de l'Académie des Sciences de Paris, Mathématiques, V. 349, N. 3-4, 219-223.
- [17] Bagdonavičius, V., Nikulin, M. (2011a). Chi-squared Goodness-of-fit Test for Right Censored Data. International Journal of Applied Mathematics and Statistics, vol. 24, p 30-50.
- [18] Bahadur Sankhya, R. R., (1958) Examples of Inconsistency of Maximum Likelihood Estimates, The Indian Journal of Statistics (1933-1960), V 20, No. 3/4 , p 207-210.
- [19] Birnbaum, Z. W. and Saunders, S. C. (1969a). A new family of life distribution, Journal of Applied Probability, vol. 6, p 319-327.

- [20] Birnbaum, Z. W. and Saunders, S. C. (1969b). Estimation for a family of life distributions with applications to fatigue, *Journal of Applied Probability*, vol. 6, p 328-347.
- [21] Bolard P, Quantin C, Esteve J, Faivre J, Abrahamowicz M. (2001a) Time dependent hazard ratio in relative survival with application in colon cancer. *Journal of Clinical Epidemiology* 54, p 986-996.
- [22] Brian, S.(2007) Méthodes numériques de résolution d'équation différentielles, Université de Provence, Institut Fresnel, Case 161 Faculté de St Jérôme, Marseille, France.
- [23] Cai, T., Huang, J., Tian, L., (2009) Regularized Estimation for the Accelerated Failure Time Model, *Biometrics*, V 65, Issue 2, 394-404.
- [24] Chen, J. (1975) Goodness of Fit Tests Under Random Censorship, PhD dissertation, Oregon State University.
- [25] Chernoff, H., Lehmann, E. L. (1954) The use of maximum likelihood estimates in tests for goodness of fit, *Ann. Math. Statist.*, 25 , 579-586.
- [26] Cox DR. (1972) Regression models and life table (with discussion). *Journal of the Royal Statistical Society, Series B* 34, p 187-202.
- [27] Cramer, U. (1946). *Mathematical Methods of statistics*, Princeton University Press, New York.
- [28] Dennis, Jr. J. E., Robert, B. S., (1983) *Numerical Methods for Unconstrained Optimization and Nonlinear Equations*, classics in applied mathematic, prentice-Hall.
- [29] Drapella A, (1993) Complementary Weibull distribution : unknown or just forgotten, *Qual Reliab Eng Int* 9 : 383-385.
- [30] Drost, F. (1988) Asymptotics for Generalized Chi-squared Goodness-of-fit Tests, amsterdam : Centre for Mathematics and Computer Sciences, CWI Tracs, 48.
- [31] Dzhaparidze, K. O. , Nikulin, M. S. (1974) On a modification of the standard statistic of Pearson, *Theory of Probability and Its Applications*, 19, 851-852.
- [32] Dzhaparidze, K. O. and Nikulin, M. S. (1992). On calculating Chi-squared type statistics, *Problems of the theory of probability distributions*, 12, 117, Nauka, St. Petersburg, Russia.
- [33] Efron, B. (1988) Logistic Regression, Survival Analysis, and the Kaplan-Meier Curve. *Journal of the American Statistical Association* 83, 414-425.

- [34] Fletcher, R. (2005) On the Barzilai-Borwein Method. Optimization and Control with Applications Applied Optimization, 96, 235-256.
- [35] Felipe R. S. de Gusmão (2011) · Edwin M., Ortega Gauss,M., Cordeiro, M. The generalized inverse Weibull distribution. Statistical Papers, 52, Issue 3, 591-619.
- [36] Fisher, R. (1928) On a property connecting the Chi-square measure of discrepancy with the method of maximum likelihood. Atti de Congresso Internazionale di Mathematici, Bologna, 6, 95-100.
- [37] Goual, H. and Seddik Ameur, N. (2014). Chi-squared type test for the AFT-generalized inverse Weibull distribution. Communication in Statistics-Theory and Method, 43, 13, 2605-2617.
- [38] Gray, R.J., and Pierce, D.A. (1985). Goodness of fit tests for censored survival data. Ann. Statist, 13, 552-563.
- [39] Greenwood, P. S., Nikulin, M. (1996) A guide to Chi-squared Testing, John Wiley and Sons, New York.
- [40] Habib, M. G. H. (1981) A Chi-Square Goodness-of Fit Test for Censored Data, PhD dissertation, Oregon State University.
- [41] Habib, M.G., Thomas, D.R. (1986) Chi-squared Goodness-of-Fit Tests For Randomly Censored Data, Annals of Statistics, V.14, N 2, p 759-765.
- [42] Haghghi, F. (2004) Fiabilité et Analyse de Survie : Le Modèle semi-paramétrique de Dégradation. Une généralisation de la famille Weibull et ses applications, Thèse de Doctorat, Université Victor Segalen Bordeaux 2.
- [43] Hahn, G. J, Shapiro, S. S. (1967) Statistical Models in Engineering, j. Wiley. N. Y, USA
- [44] Hakulinen T, Tenkanen L. (1987) Regression analysis of relative survival rates. Applied Statistics 36, p 309-317.
- [45] Haruo, Y., Kei, T., Yoshio, T.,(2011) Generalized Inverse Matrices, Statistics for Social and Behavioral Sciences, p 55-86.
- [46] Hasselman, B. (2009) nleqslv : Solve Systems of Non Linear Equations. R package version 1.4, URL <http://CRAN.Rproject.org/package=nleqslv>.
- [47] Hjort, N.L. (1990) Goodness of Fit Test in Models for Life History Data Based on Cumulative Hazard Rates, The annals of statistics, V.18, N3, p 1221-1258.

- [48] Hsuan, T.A, Robson, D.S. (1976) The χ^2 – Goodness-of-fit Tests with Moment Type Estimator, Communications in Statistics - Theory and Methods, 16 , 1509-1519.
- [49] Huan-Jyh, S., Elsayed, E. A., James, T. L., (1999) A general model for accelerated life testing with time-dependent covariates, Naval Research Logistics (NRL), V 46, Issue 3, p 303-321.
- [50] Huber-Carol, C., Balakrishnan, N., Nikulin, M.S., Mesbah, M., (2002) Goodness-Of-Fit Tests and Model Validity, Library of congres Washington USA.
- [51] Huber, C., Limnios, N., Mesbah M., Nikulin, M. (Eds) (2008) Mathematical Methods in Survival analysis, Reliability and Quality of Life, Wiley-ISTE.
- [52] Ian J., (2005) Accelerated Failure-time Models, Encyclopedia of Biostatistics, Wiley Online Library.
- [53] Ionescu, D.C, Lmmnios, N. (Eds) (1999) Statistical and Probabilistic Models in Reliability, Birkhauser Boston.
- [54] Jianmin, Z., Zhengyi J., Taosen L., Daoguo, Y., Yun-Hae, K. (2011) Parameter Estimations of Inverse Weibull Distribution, Advanced Materials Research, Volumes 199 - 200, p 564-568.
- [55] Jisheng, K., Yitian, L., Xiuhua, W.,(2006) Efficient continuation Newton-like method for solving systems of non-linear equations. Applied Mathematics and Computation 174 :2, p 846-853.
- [56] Kalbleisch JD, Prentice R. (1980) The statistical analysis of failure time data. Wiley, New York.
- [57] Kaplan EL, Meier P. (1958) Non parametric estimation from incomplete observations. Journal of the American Statistical Association 53, P 457-481.
- [58] Keller AZ, Kamath AR (1982) Reliability analysis of CNC machine tools. Reliab Eng 3 : 449-473
- [59] Kiefer, J., Wolfowitz, J. (1956) Consistency of the Maximum Likelihood Estimator in the Presence of Infinitely Many Incidental Parameters, The Annals of Mathematical Statistics, V 27, No. 4 , p 887-906.
- [60] Kim, J.H. (1993) Chi-squared Goodness-of-Fit Tests For Randomly Censored Data, Annals of Statistics, the annals of statistics, V.21, N^o3, p 1621-1639.

- [61] Lawless, J.F. (2003) Statistical Models and Methods for Lifetime Data, 2nd ed., New York : John Wiley.
- [62] LeCam, L. (1956) Locally asymptotically normal families of distributions, In University of California, Publication in Statistics , vol.3, p 37-98.
- [63] LeCam, L., Mahan,C., Singh, A. 1983. An extension of a Theorem of H.Chernoff and E.L.Lehmann. In : Recent advances in statistics, Academic Press, Orlando, p 303-332.
- [64] Lemeshko, B.Yu., Lemeshko, S.B.(2009) Distribution models for nonparametric tests for fit in verifying complicated hypotheses and maximum-likelihood estimators. Part 1. Measurement Techniques, V. 52, N 6, p 555-565.
- [65] Ludwig, F., Heinz, K., (1985) Consistency and Asymptotic Normality of the Maximum Likelihood Estimator in Generalized Linear Models, The Annals of Statistics, V 13, No. 1, p 342-368.
- [66] Meeker, W.Q., Escobar, L.A. (1998) Statistical Methods for reliability Data, John Wiley and Sons, INC.
- [67] Mirvaliev, M. (2001) An investigation of generalized Chi-squared type statistics, Ph.D. Thesis, Academy of Science, Republic of Uzbekistan, Tashkent.
- [68] Moore, D. S.n Spruill, M. C. (1975) Unified large-sample theory of general Chi-squared statistics for tests of fit, Annals of Statistics, 3, p 599-616.
- [69] M. Shuaib Khan, G.R Pasha, A. Hesham Pasha, (2008) Theoretical Analysis of Inverse Weibull Distribution, WSEAS TRANSACTIONS on MATHEMATICS, Issue 2, Volume 7, p 30-38.
- [70] Mudholkar GS, Kollia GD (1994) Generalized Weibull family : a structural analysis, Commun Stat Ser A 23 : 1149-1171.
- [71] Mudholkar G.S, Srivastava D.K, Kollia G.D (1996) A generalization of the Weibull distribution with application to the analysis of survival data, J Am Stat Assoc, 91, p 1575-1583.
- [72] Nelson, W. (1990) Accelerated Testing : Statistical Models, Test Plans, and Data Analysis, John Wiley and Sons : New York .
- [73] Newby, M.,(1981) Accelerated failure time models for reliability data analysis, Reliability Engineering et System Safety, V 20, Issue 3, P 187-197.

- [74] Nikulin, M.S., Haghghi, F. (2004) A Chi-Squared Test for the Generalized Power Weibull Family for the Head-and-Neck Cancer Censored Data, journal of Mathematical Sciences Volume 133, Number 3, p 1333-1341.
- [75] Nikulin, M. Haghghi, F. (2009) On the power generalized Weibull family : model for cancer censored data, METRON-International Journal of Statistics, vol. LXVII, n.1, p 75-86.
- [76] Nikulin, M.S., Limnios, N., Balakrishnan, N., Kahle, W., Huber, C. (Eds.) (2009) Advances in Degradation Modeling : Applications to Reliability, Survival Analysis and Finance, Birkhäuser Boston, 454p.
- [77] Nikulin, M.S. (1973a) Chi-square Test For Continuous Distributions with Shift and Scale Parameters, teor. Veroyatn. Primen., 18, No. 3, p 559-568.
- [78] Nikulin, M. S. (1973b) On a Chi-square test for continuous distributions, Theory of Probability and its Applications, 18, p 638-639.
- [79] Nikulin. M.s, (1973c). On a Chi-squared test for continuous distributions, Theory of Probability and its Applications. vol.18, 3, p 638-639.
- [80] Nikulin, M.S., Solev, V.N. (1999) Chi-squares Goodness-of-fit Test for Doubly Censored Data with Applications in Survival Analysis and Reliability, In : Statistical and Probabilistic Models in Reliability, D.C.Ionescu, N.Limnios (eds), Birkhauser, Boston, p 101-112.
- [81] Nikulin, M.S., Gerville-Réache, L., Couallier, V. (2007) Statistiques des Essais Accélérés, Hermès : London.
- [82] Pawlas, P., Szyna, D., (2000) Characterizations of the inverse Weibull distribution and generalized extreme value distributions by moments of Kth record values, Applicationes mathematicae, 27,2 , p 197-202.
- [83] Pena, E. (1998a). Smooth goodness-of-fit tests for composite hypothesis in hazard based models, Ann. Statist. 26, p 1935-1971.
- [84] Pena, Edsel, A. (1998b). Smooth goodness-of-fit tests for the baseline hazard in Cox's proportional hazards model. J. Am. Stat. Assoc., 93, No.442, p 673-692.
- [85] Penrose,R., (1955) A generalized inverse for matrices, Philosophical Society Mathematical Proceedings of the Cambridge Philosophical Society, V 51, Issue 03, p 406-413.

- [86] Peter K. Kitanidis, Robert W. Lane, (1985) Maximum likelihood parameter estimation of hydrologic spatial processes by the Gauss-Newton method, *Journal of Hydrology*, V 79, Issues 1-2, P 53-71.
- [87] Rao, K.C., Robson, D.S. (1974) A chi-square statistic for goodness-of-fit tests within the exponential family, *Communications in Statistics*, 3, p 1139-1153.
- [88] Ravi, V., Paul, D. Gilbert. (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*, V 32, Issue 4.
- [89] Rykov, V.V., Balakrishnan, N., Nikulin, M.S. (Eds.) (2010) Mathematical and Statistical Models and Methods in Reliability, Applications to Medicine, Finance, and Quality Control Series : Statistics for Industry and Technology, Birkhäuser Boston, 458 p.
- [90] Saaidia, N., Tahir, R., Seddik-Ameur, N. (2010) The Inverse Gaussian Distribution and its Applications in Accelerated Life Testing when the Hazard Rate Function has cup shape, International Conference on Accelerated Life Testing, Reliability-based Analysis and Design 19-21 Mai 2010, Clermont-Ferrand.
- [91] Sandro, I., Valerio, P., Francesco, Z., (1979) A New Method for Solving Nonlinear Simultaneous Equations. *SIAM Journal on Numerical Analysis* 16 :5, p 779-789.
- [92] Saunders, S.C., Lawless , J.F. (2007) Reliability, life testing and the prediction of service lives : for engineers and Scientists, Springer New York, 310 p.
- [93] Seal, H.L. (1969) Stochastic Theory of a Risk Business, New York :John Wiley & Sons Inc
- [94] Sedyakin, N.M. (1966). On one physical principle in reliability theory, *Techn. Cybernet-ics*, 3, 80-87.
- [95] Small, CG., Wang, J. (2003) Numerical Methods for Nonlinear Estimating Equations. Oxford University Press, New York.
- [96] Van Der Vaart, A.W., (1998) Asymptotic Statistics. Cambridge Series in Statistics and probabilistic Mathematics, Cambridge : Cambridge University Press.

- [97] Voinov, V., Alloyarova, R., Pya, N. (2008) A Modified Chi-squared Goodness-of-fit Test for the Three parameter Weibull Distribution and its Applications in Reliability, In : Mathematical Methods in Survival Analysis, Reliability and Quality of Life (Eds. Huber C., Limnios N., Mesbah M., Nikulin M.), Wily-ISTE.
- [98] Voinov, V., Alloyarova, R., Pya, N. (2008) Recent Achievemets in Modified Chi-squared Goodness-of-fit Testing, In :Statistical Models and Methods for Biomedical and Technical Systems,(Eds. F. Vonta, M. Nikulin, N. Limnios, C. Huber.), Birkhäuser, Boston, p 241-258.
- [99] Voinov, V., Pya, N., Alloyarova, R. (2009) A Comparative Study of Some Modified Chi-squared Tests, Communi. in Stat.-Simul. and Comput., 38 :2, p 355-367.
- [100] Voinov, V., Nikulin, M. (1993) Unbiased Estimators and Their Applications, V.1 Univariate case, Dordrecht : Kluwer Academic Publishers.
- [101] Voinov, V., Nikulin, M., Balakrishnan, N.(2013). Chi-Squared Goodness of Fit Tests with Applications. Elsevier : Amesterdam, Bosten, etc.
- [102] Von Alven, W. H. (ed.) (1964) Reliability Engineering by ARINC, Englewood Cliffs, N.J.Prentice-Hall, Inc.
- [103] Wei LJ. (1984) Testing goodness of fit for proportional hazard model with censored observations. Journal of the American Statistical Association 79, p 649-652.
- [104] Xiao, Wang, Dihua, Xu. (2010) An Inverse Gaussian Process Model for Degradation Data, Technometrics, V. 52, No 2, p 188-197.
- [105] Xiaojun, C., Tetsuro Y.,(1989) Numerical Functional Analysis and Optimization, Convergence domains of certain iterative methods for solving nonlinear equations, Mathematics of Computation, V 10, Issue 1-2, p 37-48.
- [106] Zhang, B. (1999) A Chi-squared goodness-of-fit test for logistic regression models based on case-control data, Biometrika, 86, p 531-539.
- [107] Zucker DM, Karr AF. (1990) Nonparametric survival analysis with time-dependent covariate effects : a penalized partial likelihood approach. The Annals of Statistics 18, p 329-353.