



الجمهورية الجزائرية الديمقراطية الشعبية
République Algérienne Démocratique et Populaire
وزارة التعليم العالي والبحث العلمي
Ministère de l'Enseignement Supérieure et de la Recherche Scientifique

جامعة وهران 2 محمد بن أحمد
Université d'Oran 2 Mohamed Ben Ahmed
معهد الصيانة و الأمن الصناعي
Institut de Maintenance et de Sécurité Industrielle

Département de Maintenance en Electromécanique

MÉMOIRE

Pour l'obtention du diplôme de Master

Filière : Génie Industriel

Spécialité : Maintenance-Fiabilité-Qualité

Thème

Conception d'une Toolbox de fiabilité sous Matlab

Présenté et soutenu publiquement par :

BEN ALI Amer Elmouataz et BOUAFIA Djillali Chakib

Devant le jury composé de :

Nom et Prénom	Grade	Etablissement	Qualité
BOUHAHS Mohamed	MCB	Université d'Oran2	Président
YAHIAOUI Reda	MCB	Université d'Oran2	Encadreur
NOUREDDINE Rachid	Pr	Université d'Oran2	Examineur

Année 2019/2020

Remerciements et Dédicaces

Nous tenons à exprimer nos remerciements à MR. YAHIAOUI Reda pour nous avoir encadré dans la réalisation de ce travail et pour son aide qu'il nous a apporté ainsi que les membres du jury.

Nous remercions aussi tous nos amis pour leur soutien moral et toutes personnes ayant aidé de près ou de loin pour réaliser ce travail.

Ma mère la source et l'espoir. Mon père le repère et l'exemplaire. A vous je dédie ce modeste mémoire pour m'avoir encouragé et poussé à atteindre l'idéal, m'avoir soutenu tout au long de mes études.

Je dédie ce modeste travail à Mes frères, toute ma famille, tous mes fidèles amis, Je dédie enfin ce mémoire à toute personne ayant contribué de près ou de loin à sa concrétisation.

Nomenclature

$R(t)$: fonction de fiabilité.

$F(t)$: fonction de répartition (densité cumulative).

$f(t)$: fonction de densité.

$z(t)$: Taux défaillance (Failure rate).

$h(t), \lambda(t)$: Taux défaillance (Hazard rate function).

$H(t)$: Taux défaillance cumulé (Cumulative Hazard rate function).

$M(t)$: Fonction d'échelon.

$m(t)$: taux d'occurrence des défaillances (TOD).

$\bar{M}(t)$: Estimation de $M(t)$.

$E[T]$: L'espérance mathématique.

$V(T)$: La variance mathématique.

$L(t)$: La durée de vie résiduelle moyenne.

A : La Disponibilité.

β : Le paramètre de forme (Loi Weibull).

γ : Le paramètre de position (Loi Weibull).

η : Le paramètre de position (Loi Weibull).

μ : paramètre d'échelle (Loi Normal).

σ : paramètre de forme (Loi Normal).

μ : paramètre de position (Loi GEV).

σ : paramètre d'échelle (Loi GEV).

k : paramètre de forme (Loi GEV).

$\Gamma(x)$: la fonction intégrale Eulérienne de seconde espèce.

$\Phi(z)$: la fonction de répartition (la distribution cumulative normale standard).

$\Gamma(\theta)$: la fonction gamma.

U : Statistique de Khi-deux.

A^2 : Statistique d'Anderson Darling.

A^* : La distribution asymptotique par Stephens.

D_n : règle de décision de Kolmogorov Smirnov.

$K_n(y)$: Fonction de Glivenko et Kolmogorov.

$n\omega^2$: Statistique de test de Cramer-Von-Mises.

Acronymes

FEI: la fiabilité en tant qu'ingénierie.

FES: la fiabilité en tant que services.

MTTF: Mean time to failure.

MTBF: Mean time between failures.

MTBR: Mean time between replacements.

MTTR: Mean time to repair.

LUR : line-replaceable unit.

PDF: Probability density function.

CDF: Cumulative density function.

TOD : taux d'occurrence des défaillances.

GUI : Graphical User Interface.

MRL : Mean résiduel life.

Listes des figures

Figure I.1 : Le rôle de la fiabilité dans le cycle de vie d'un produit.

Figure I.2 : Courbe du taux de risque en baignoire.

Figure I.3 : Scénarios de remplacement pour un système réparable.

Figure I.4 : Le MTTR du remplacement d'un moteur d'avion.

Figure II.1 : Le taux de défaillance de la fonction de Weibull.

Figure II.2 : Fonction de fiabilité de Weibull.

Figure II.3 : Fonction de densité de probabilité de Weibull.

Figure II.4 : Fonction de taux de défaillance instantané de la distribution de Weibull.

Figure II.5 : la PDF normale pour différentes moyennes et variances.

Figure II.6 : Représentation de la FDP de loi Log-normal.

Figure II.7 : Représentation de la PDF gamma pour différentes valeur de λ et θ .

Figure II.8 : tableau de niveau de signification.

Figure III.1 : L'outil GUIDE de Matlab.

Figure III.2 : L'interface graphique de la toolbox de fiabilité.

Figure III.3 : Exemple de représentation des résultats du test d'ajustement.

Figure III.4 : exemple de résultats comparables du test d'ajustement

Figure III.5 : Interface graphique pour le calcul de la MRL

Figure III.6 : Introduction des données.

Figure III.7 : Représentation des fonctions de survie

Figure III.8 : Représentation des fonctions de fiabilité non paramétrique par noyau (Kernel).

Figure III.9-a : Résultats du test d'ajustement d'Anderson-Darling.

Figure III.9-b : Résultats du test d'ajustement de Khi-deux.

Figure III.10 : Interface graphique pour la représentation des fonctions de fiabilité.

Figure III.11-a : Représentation graphique de la PDF.

Figure III.11-b : Représentation graphique de fonction de survie.

Figure III.11-c : Représentation graphique de la fonction CDF.

Figure III.11-d : Représentation graphique de fonction de Hazard cumulative.

Figure III.12 : Représentation de la durée de vie résiduelle.

Sommaire

Remerciements et Dédicaces	1
Nomenclature	2
Acronymes	4
Listes des figures	5
Sommaire	7
Introduction générale	9
Chapitre I : Maintenance et fiabilité	10
I.1 Introduction	11
I.2 Définition de la fiabilité et du taux de défaillance	11
I.2.1 La fiabilité et les mesures probabilistes	12
I.2.2 Fonction de taux de défaillance et de taux de risque	13
I.2.3 Courbe en baignoire du taux de risque	15
I.2.4 Taux d'intensité des défaillances	15
I.3 Durée de vie moyenne et durée de vie résiduelle moyenne	16
I.3.1 Le temps moyen de fonctionnement avant panne (Mean-Time-to-Failure)	16
I.3.2 Temps moyen entre pannes (Mean-Time-Between-Failures)	17
I.3.3 Le temps moyen entre réparations (Mean-Time-Between-Replacements)	17
I.3.4 Durée de vie résiduelle moyenne	19
I.4 Temps d'arrêt et disponibilité du système	19
I.4.1 Temps moyen de réparation (Mean-Time-to-Repair)	19
I.4.2 Disponibilité du système	20
I.5 Conclusion	20
Chapitre II : Lois de fiabilité et tests d'ajustement	21
II.1 Introduction	22
II.2 Méthodes paramétrique	22
II.2.1 Loi exponentielle	22

II.2.2 Loi de Weibull	23
II.2.3 Loi normale	27
II.2.4 Loi log-normale	28
II.2.5 Loi Gamma	29
II.2.6 Loi d'extremum généralisée	31
II.3 Méthodes non paramétrique	32
II.3.1 Estimation d'une fonction de densité de probabilité par noyaux	32
II.3.2 Méthode d'estimation de Kaplan-Meier	33
II.4 Test d'hypothèse	35
II.4.1 Le Test du Khi-Deux (khi square χ^2)	35
II.4.2 Test d'Anderson-Darling	36
II.4.3 Test de Kolmogorov-Smirnov (1933, 1939)	38
II.4.4 Test de Cramer-Von-Mises	39
II.4.5 Tests de normalité et d'exponentialité	39
II.5 Conclusion	40
Chapitre III : Présentation de la toolbox et application	41
III.1 Introduction	42
III.2 Toolbox en Matlab	42
III.2.1 Les interfaces graphiques GUI (Graphical User Interface)	42
III.2.2 L'outil GUIDE	42
III.3 Présentation de la toolbox de fiabilité	43
III.3.1 Menu de la toolbox	44
III.3.2 Interface graphique des résultats	45
III.4 Application de la toolbox dans le cas de l'usure des outils de coupe	46
III.4.1 Analyse de fiabilité	47
III.4.2 Analyse non paramétrique	48
III.4.3 Analyse paramétrique	49
III.5 Conclusion	53
Conclusion générale	54
Références bibliographiques	55

Introduction générale

L'analyse de la fiabilité constitue une phase indispensable dans toute étude de sûreté de fonctionnement. A l'origine, la fiabilité concernait les systèmes à haute technologie (centrales nucléaires, aérospatial). Aujourd'hui, la fiabilité est devenue un paramètre clé de la qualité et d'aide à la décision, dans l'étude de la plupart des composants, produits et processus "grand public": Transport, énergie, bâtiments, composants électroniques, composants mécaniques, etc.

L'analyse de fiabilité est un domaine spécialisé dans les statistiques mathématiques et est développée pour traiter le type spécial de variables aléatoires temporelles (durée de vie, temps de défaillance, temps de survie, etc.)

Le délai avant défaillance ou la durée de vie d'un article est intimement lié à sa fiabilité, et c'est une caractéristique qui varie d'un système à l'autre, même s'ils sont identiques dans la conception et la structure. Par exemple, si nous utilisons les mêmes de composants automobiles dans différentes automobiles et observons leur défaillance individuelle On ne s'attendrait pas à ce qu'ils connaissent tous les mêmes périodes d'échec. Les temps d'échec pour les composants utilisés dans les différentes automobiles seraient différents et seraient définis par une variable aléatoire. Le comportement de la variable aléatoire peut être modélisé par une distribution de probabilité, qui est une description mathématique d'un phénomène aléatoire qui consiste en un espace d'échantillonnage et une façon d'attribuer des probabilités aux événements. La base de l'analyse de fiabilité consiste à modéliser la durée de vie par une distribution de probabilité appropriée et de caractériser le comportement de la vie à travers la distribution choisie.

Donc L'analyse de fiabilité nous permet de répondre à des questions telles que :

- Quelle est la probabilité qu'une unité tombe en panne avant un moment donné ?
- Quel est le pourcentage d'éléments qui dureront plus longtemps qu'un certain temps ?
- Quelle est la durée de vie prévue d'un élément ?

Cependant, les chercheurs et les praticiens des domaines de fiabilité et de l'informatique ont estimé qu'ils seraient judicieux que les techniques statistiques d'intérêt commun peuvent être partagées, ce qui est extrêmement bénéfique. Le présent travail est conduit dans ce sens et vise à concevoir une toolbox, sous le langage Matlab, dédiée au domaine de la maintenance.

Chapitre I : Maintenance et fiabilité

I.1 Introduction

La fiabilité est une science relativement récente dont on peut situer approximativement les débuts vers les années 1960. Elle s’est développée très rapidement, elle a des applications dans de nombreux domaines. Comme définition, la fiabilité est une approche statistique visant à décrire la sécurité et la capacité d'un système ou d'un composant à fonctionner dans des conditions données pendant une période de temps déterminée en présence d'une incertitude.

I.2 Définition de la fiabilité et du taux de défaillance

L'ingénierie de la fiabilité est un domaine interdisciplinaire qui étudie, évalue et gère les performances des composants et des systèmes tout au long de leur durée de vie, tels que les automobiles, les éoliennes, avions, Internet, appareils médicaux, système d'alimentation électrique et radars, etc. autres [1]. Ces systèmes et les équipements sont largement utilisés dans les secteurs du commerce et de la défense, allant de la fabrication, l'énergie, les transports, les soins de santé, les communications et les opérations militaires.

Le cycle de vie d'un produit comprend généralement cinq phases: conception/développement, introduction du produit, l'expédition de volumes, la saturation du marché et le retrait progressif. La **Figure I.1** illustre l'interdépendance des cinq phases. La fiabilité joue un double rôle tout au long du cycle de vie d'un produit : la fiabilité en tant qu'ingénierie (FEI) et la fiabilité en tant que services (FES). FEI englobe la conception de la fiabilité, la planification de la croissance de la fiabilité, ainsi que la garantie et la maintenance. FES se concentre sur la planification et la gestion d'un système d'inventaire des produits réparables, la fourniture de pièces de rechange, et le recyclage et la refabrication des produits en fin de vie. La FEI et la FES ont été étudiées de manière intensive, mais souvent séparément en ingénierie de la fiabilité et les communautés de gestion des

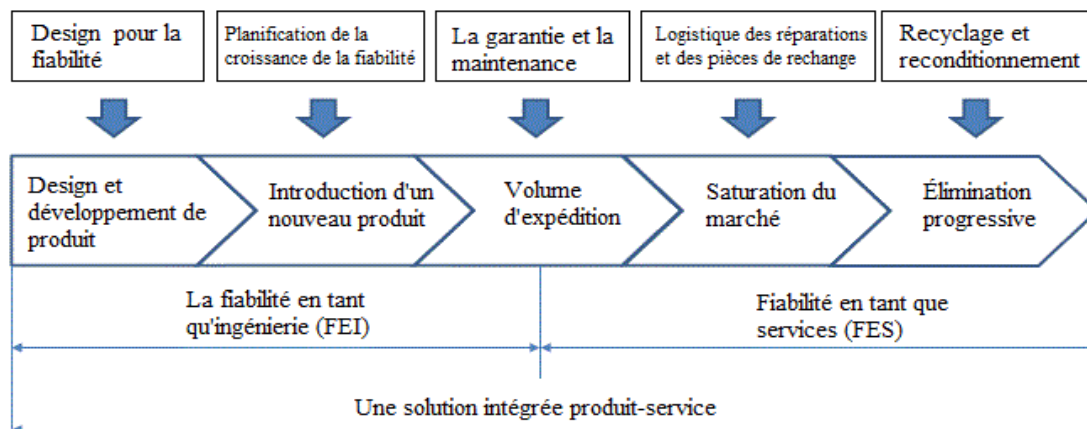


Figure I.1 Le rôle de la fiabilité dans le cycle de vie d'un produit.

opérations. La fusion de la FEI et de la FES est motivée principalement par la concurrence mondiale intense, la compression du cycle de conception des produits, la volatilité de la chaîne d'approvisionnement, la durabilité environnementale et l'évolution des besoins des clients. Il y a une tendance croissante selon laquelle la FEI et la FES seront intégrées de manière transparente dans le cadre du système produit-service, qui offre une solution de fiabilité groupée aux clients.

Les ingénieurs de fiabilité doivent travailler avec les ingénieurs du matériel et de logiciel, le groupe d'achat des composants, le département de fabrication et d'exploitation, les techniciens de soutien et de réparation sur le terrain, les planificateurs de la logistique et des stocks, et l'équipe de marketing pour identifier et éliminer les principales causes fondamentales en temps utile, mais de manière rentable. Par conséquent, un ingénieur de fiabilité doit posséder un large éventail de compétences allant de l'ingénierie, de la physique, des mathématiques, des statistiques et de la recherche opérationnelle à la gestion d'entreprise. Enfin, et ce n'est pas le moins important, un ingénieur en fiabilité doit posséder une forte capacité de communication afin de diriger des initiatives de mesures correctives, de résoudre des objectifs contradictoires entre différentes unités organisationnelles et d'apporter une contribution précieuse à la conception des produits, à la production en volume et au service après-vente.

I.2.1 La fiabilité et les mesures probabilistes

La fiabilité est définie comme la capacité d'un système ou d'un composant à fonctionner de la manière requise dans des conditions précises pendant une période de temps déterminée [2] [3]. Elle est souvent mesurée comme une probabilité de défaillance ou une possibilité de disponibilité. Soit T une variable aléatoire non négative représentant la durée de vie d'un système ou d'un composant, la fonction de fiabilité, désignée par $R(t)$, est alors exprimée par

$$R(t) = P\{T > t\}, \text{ pour } t = 0 \quad (\text{I.1})$$

C'est la probabilité que T dépasse une durée de vie attendue t qui est généralement spécifiée par le fabricant ou le client. Les éléments clés de la définition de l'équation (I.1) sont mis en évidence ci-dessous.

- La fiabilité est prédite sur la base de la "fonction prévue" ou du "fonctionnement" sans défaillance. Cependant, si les pièces individuelles sont bonnes mais que le système dans son ensemble n'atteint pas les performances prévues, il est alors toujours classé comme une défaillance. Par exemple, un système solaire photovoltaïque ne produit pas d'électricité pendant la nuit. Par conséquent, la fiabilité de l'approvisionnement en énergie est nulle même si les panneaux solaires et les onduleurs DC-AC sont bons.

- La fiabilité est limitée à l'exploitation dans des conditions explicitement définies. Il est pratiquement impossible de concevoir un système pour des conditions illimitées. Un véhicule électrique doit fonctionner sous des conditions beaucoup plus sévères que celles d'une voiture de golf, même si elles sont alimentées par le même type de batterie. Les conditions de fonctionnement

et l'environnement doivent être pris en compte lors de la conception et de l'essai d'un nouveau produit.

- La fiabilité s'applique à une période de temps déterminée, ce qui signifie que tout système finira par tomber en panne. L'ingénierie de la fiabilité garantit qu'un système fonctionnera sans défaillance avant le temps t .

La relation entre la distribution des temps de défaillance $F(t)$ et la fonction de fiabilité $R(t)$ est régie par :

$$F(t) = 1 - R(t) = 1 - P\{T > t\} \quad (I.2)$$

En statistique, $F(t)$ est également appelée fonction de distribution cumulative (CDF). Soit $f(t)$ la fonction de densité de probabilité (PDF) ; la relation entre $R(t)$ et $f(t)$ est donnée comme suit:

$$R(t) = \int_t^{\infty} f(x)dx = 1 - \int_0^t f(x)dx \quad (I.3)$$

I.2.2 Fonction de taux de défaillance et de taux de risque

Étant donné qu'un système fonctionne au temps t , la probabilité que le système tombe en panne dans l'intervalle $[t, t + \Delta t]$ est

$$\begin{aligned} P\{t < T \leq t + \Delta t | T > t\} &= \frac{P\{t < T \leq t + \Delta t | T > t\}}{P\{T > t\}} = \frac{P\{t < T \leq t + \Delta t\}}{P\{T > t\}} \\ &= \frac{R(t) - R(t + \Delta t)}{R(t)} \end{aligned} \quad (I.4)$$

Le résultat est dérivé sur la base du théorème de Bayes en réalisant $P\{A, B\} = P\{A\}$, où A est le cas où le système échoue dans l'intervalle $[t, t + \Delta t]$ et B est le cas où le système survit jusqu'au temps t .

Le taux de défaillance, désigné par $z(t)$, est défini dans un intervalle de temps $[t, t + \Delta t]$ comme la probabilité qu'une défaillance par unité de temps se produisant dans cet intervalle étant donné que le système a survécu jusqu'à t . C'est-à-dire,

$$z(t) = \frac{P\{t < T \leq t + \Delta t | T > t\}}{(t + \Delta t) - t} = \frac{R(t) - R(t + \Delta t)}{\Delta t R(t)} \quad (I.5)$$

Bien que le taux de défaillance $z(t)$ dans l'équation (I.5) soit souvent considéré comme la probabilité qu'un échec se produit dans un intervalle spécifié $[t, t + \Delta t]$ sachant qu'aucun échec

n'est survenu avant le temps t , il n'est en fait pas une probabilité car $z(t)$ peut dépasser 1. Par conséquent, le taux de défaillance représente la fréquence à laquelle un système ou un composant tombe en panne et est exprimé en défaillances par unité de temps. Le taux de défaillance réel d'un produit ou d'un système est étroitement lié à l'environnement d'exploitation et à l'utilisation par le client [4].

La fonction de risque, également connue sous le nom de taux de défaillance instantané, est définie comme la limite du taux de défaillance lorsque Δt s'approche de zéro. la fonction de taux de risque $h(t)$ est obtenue comme suit :

$$\begin{aligned} h(t) &= \lim_{\Delta t \rightarrow 0} \frac{R(t) - R(t + \Delta t)}{\Delta t R(t)} = -\frac{1}{R(t)} \lim_{\Delta t \rightarrow 0} \frac{R(t + \Delta t) - R(t)}{\Delta t} \\ &= -\frac{1}{R(t)} \times \frac{dR(t)}{dt} = \frac{f(t)}{R(t)} \end{aligned} \quad (I.6)$$

L'équation (I.6) représente un résultat important car elle régit la relation entre $h(t)$, $f(t)$ et $R(t)$. Alternativement, à partir de l'équation (I.6), la fonction de fiabilité $R(t)$ peut être exprimée comme suit :

$$R(t) = \exp\left(-\int_0^t h(x)dx\right) \quad (I.7)$$

Désignons $H(t)$ comme la fonction de taux de risque cumulé, alors :

$$H(t) = \int_0^t h(x)dx \quad (I.8)$$

En substituant l'équation (I.8) à (I.7), la fonction de fiabilité $R(t)$ peut également être exprimée comme :

$$R(t) = e^{-H(t)} \quad (I.9)$$

Le taux de défaillance et le taux de risque sont souvent utilisés de manière interchangeable dans la littérature sur la fiabilité et les applications industrielles pour la modélisation des systèmes non réparables. Les deux mesures sont également applicables aux systèmes réparables dont les composants individuels ne sont pas réparables. En outre, les valeurs du taux de risque et du taux de défaillance sont toujours non négatives.

I.2.3 Courbe en baignoire du taux de risque

Comme son nom l'indique, la courbe du taux de risque des baignoires est dérivée de la forme de la section transversale d'une baignoire. Comme le montre la **Figure I.2**, la courbe des taux de risque en baignoire se compose de trois types différents de profil : les défaillances de mortalité infantile lors de l'introduction initiale du produit, le taux constant pendant sa période d'utilisation et le taux croissant de défaillances dû à l'usure ou à la dégradation continue et excessive du produit.

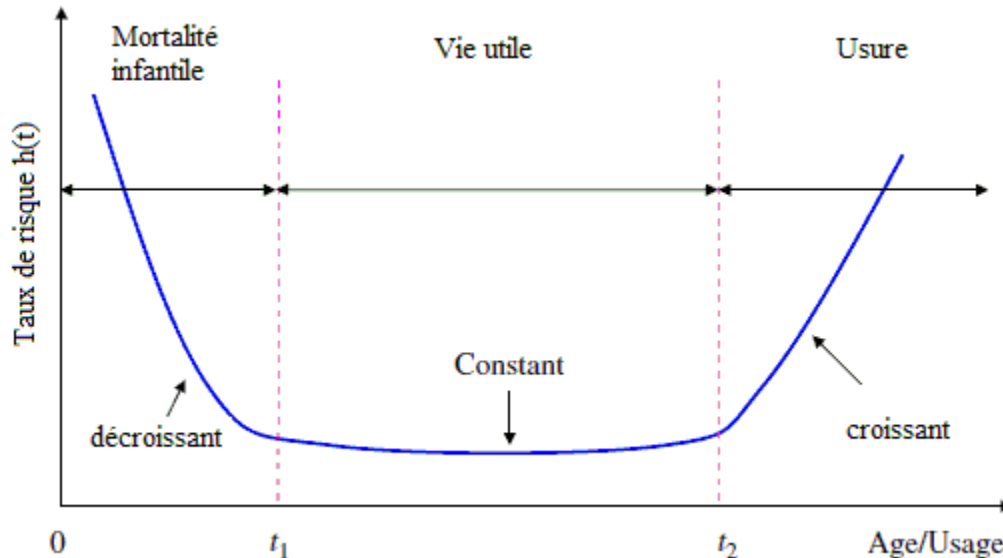


Figure I.2 Courbe du taux de risque en baignoire.

Dans les secteurs de l'électronique militaire et de l'électronique grand public, la mortalité infantile est souvent éliminée grâce à ce que l'on appelle le processus d'examen environnemental préalable, c'est-à-dire qu'avant d'être expédiés au client, les produits sont testés dans des conditions de fonctionnement plus difficiles (par exemple, température, humidité, vibrations et tension électrique) pendant une période déterminée afin de filtrer les unités faibles du parc de produits. Ce processus est adopté principalement pour les applications critiques en termes de mission ou de sécurité, car il réduit considérablement la possibilité d'apparition de défaillances du système au début de sa vie. Bien que la courbe de la baignoire soit utile, tous les produits ou systèmes ne suivent pas un profil de taux de danger de type baignoire.

I.2.4 Taux d'intensité des défaillances

En cas de défaillance, le système peut être réparé et remis en état, ce qui est fait en remplaçant les composants défectueux par d'autres neufs ou en bon état. Au fil du temps, la fréquence des défaillances peut augmenter, diminuer ou rester à un niveau constant en fonction de la politique de maintenance. Le système, une fois réparé, peut être ramené à l'une des conditions suivantes : bon comme neuf, bon comme d'habitude, et quelque part entre les deux. La fonction d'intensité de

la défaillance est une mesure généralement utilisé pour mesurer l'occurrence des défaillances par unité de temps pour un système réparable. Une distinction est faite entre le taux de risque et le taux d'intensité des défaillances. Le premier est utilisé pour caractériser le temps jusqu'à la première défaillance d'un composant, tandis que la dernière traite des défaillances récurrentes du même système. D'où, le taux d'intensité des défaillances est également appelé taux d'occurrence des défaillances (TOD). Soit $M(t)$ les défaillances (ou réparations) cumulatives survenues dans un système réparable au cours de $[0, t]$. Le TOD, dénommée $m(t)$, est défini comme suit :

$$m(t) = \frac{M(t)}{t} \quad \text{pour } t > 0 \quad (\text{I. 10})$$

L'unité de TOD est le nombre de défaillances par unité de temps (taux d'occurrence des défaillances). Il faut noter que les défaillances dans un système réparable peuvent se produire sur différents types de pièces ou sur les mêmes types de pièces, mais sur des équipements différents.

Il faut noter que $M(t)$ dans l'équation (I. 10) est une fonction d'échelon qui marque un saut vers le haut à chaque fois qu'une défaillance se produit et reste à ce niveau jusqu'à la défaillance suivante. Chaque système aura sa propre fonction $M(t)$ observée dans le temps. Si un certain nombre de courbes $M(t)$ sont observées à partir de n systèmes similaires et que les courbes sont moyennées, nous aurions une estimation de $M(t)$, dénommée $\bar{M}(t)$. C'est-à-dire,

$$\bar{M}(t) = \int_0^t m(x) dx \quad (\text{I. 11})$$

Où $m(t)$ est la fonction TOD pour le groupe de n systèmes. Le TOD est parfois aussi appelé taux de réparation, à ne pas confondre avec la durée d'exécution d'une réparation.

I.3 Durée de vie moyenne et durée de vie résiduelle moyenne

I.3.1 Le temps moyen de fonctionnement avant panne (Mean-Time-to-Failure)

Le temps moyen avant défaillance (MTTF) est une mesure quantitative couramment utilisée pour évaluer la fiabilité des systèmes ou produits non réparables. Il mesure la durée de vie prévue d'un composant ou d'un système avant qu'il ne tombe en panne. Par exemple, un panneau solaire photovoltaïque est considéré comme un système non réparable avec un MTTF typique entre 20 et 30 ans. La MTTF d'un pneu pour véhicules commerciaux varie entre 48 000 km et 96 000 km. Comme ces articles ne sont pas réparables en cas de défaillance, ils sont soit mis au rebut, soit recyclés à des fins de protection de l'environnement. Dans l'industrie, les produits non réparables sont également appelés articles consommables. Soit n le nombre de systèmes non réparables opérant sur le terrain et $t_1, t_2, t_3, \dots, t_n$ le temps de défaillance correspondant à chaque système, le MTTF est défini par :

$$MTTF = \frac{1}{n} \sum_{i=1}^n t_i \quad (I.12)$$

Si la taille de l'échantillon n est suffisamment importante, la distribution du temps T avant défaillance peut être déduite statistiquement. Le MTTF est alors équivalent à la valeur attendue de T , à savoir :

$$MTTF = E(T) = \int_0^{\infty} t f(t) dt \quad (I.13)$$

Où $f(t)$ est le PDF de la vie du système. Le MTTF peut également être exprimé comme l'intégration de $R(t)$ sur $[0, +\infty)$ en effectuant l'intégration par partie dans Eq. (I.13). Il en résulte :

$$MTTF = \int_0^{\infty} t f(t) dt = \int_0^{\infty} R(t) dt \quad (I.14)$$

I.3.2 Temps moyen entre pannes (Mean-Time-Between-Failures)

Pour un système réparable, le temps moyen entre deux défaillances consécutives est caractérisé par le temps moyen entre défaillances (MTBF). Le MTBF et le MTTF mesurent tous les deux le temps moyen de fonctionnement d'un système, mais le MTBF est utilisé pour les systèmes réparables, par opposition au MTTF qui n'est que pour les systèmes non réparables.

Par exemple, les pneus sont considérés comme un composant non réparable, et leur fiabilité est caractérisée par le MTTF. Cependant, une voiture équipée de quatre pneus est un système réparable, et sa fiabilité est mesurée par le MTBF. Soit $t_1, t_2, t_3, \dots, t_n$ les temps entre deux défaillances consécutives, le MTBF d'un système réparable est donné par :

$$MTBF = \frac{1}{n} \sum_{i=1}^n t_i \quad (I.15)$$

Lors de l'estimation du MTBF, il convient d'exclure les temps d'arrêt liés à l'attente des techniciens de réparation, le temps d'expédition des pièces de rechange, les diagnostics de défaillance et les retards administratifs. En d'autres termes, le MTBF ne mesure que le temps de fonctionnement moyen lorsqu'un système réparable est disponible pour la production pendant deux pannes consécutives.

I.3.3 Le temps moyen entre réparations (Mean-Time-Between-Replacements)

Pour un système réparable composé de plusieurs types de composants, le temps moyen entre les remplacements (MTBR) est une mesure de fiabilité associée à un type de composant spécifique. Les composants de différents types peuvent être classés en unités réparables ou non réparables.

Par exemple, un train d'atterrissage d'avion est réparable tandis que les pneus sont traités comme non réparables. Si les composants sont réparables, ils sont également appelés unité remplaçable en ligne (LRU). Si les composants ne sont pas réparables, ils sont appelés une pièce consommable. Par exemple, les Éolienne modernes sont un système réparable, et chaque turbine comprend généralement trois pales, un palier principal, une boîte de vitesses, un générateur, une électronique de puissance et d'autres unités de commande. Les composants tels que la boîte de vitesses et le générateur sont des LRU car ce sont des unités réparables. Un générateur défectueux après avoir été réparé dans l'atelier de réparation peut être réutilisé dans d'autres Éolienne. Les aubes et les roulements de turbine sont traités comme des pièces consommables. En cas de panne, ils sont jetés ou recyclés au lieu d'être réparés et réutilisés.

Il existe deux types de remplacement selon que le composant a échoué soudainement ou a atteint l'âge de la maintenance programmée (mais n'a pas échoué). Ce dernier type est appelé remplacement préventif. Mathématiquement, le MTBR signifie le temps moyen entre deux remplacements consécutifs qui tient compte à la fois des remplacements en cas d'échec et les remplacements prévus. Soit t_1, t_2, \dots, t_n les temps correspondants aux remplacement effectués et soit $\tau_1, \tau_2, \dots, \tau_m$ les temps prévus de remplacement. En se référant aux scénarios de remplacement dans la **Figure I.3**, le MTBR d'un type de composant particulier dans un système réparable est donné par :

$$MTBR = \frac{1}{m + n} \left(\sum_{i=1}^m \tau_i + \sum_{i=1}^n t_i \right) \quad (I.16)$$

En maintenance préventive, l'intervalle de remplacement est souvent programmé à l'avance avec une durée fixe. Si le délai de remplacement pratique est court et peut être ignorée, la valeur attendue du MTBR dans le cadre de la politique d'intervalle de remplacement constant peut être obtenue par :

$$MTBR = \tau R(\tau) + \int_0^{\tau} t f(t) dt = \int_0^{\tau} R(t) dt \quad (I.17)$$

Il faut noter que $\tau R(\tau)$ saisit tous les événements de remplacement programmés à intervalle fixe τ et $\int_0^{\tau} t f(t) dt$ représente les remplacements de pannes survenant avant τ .

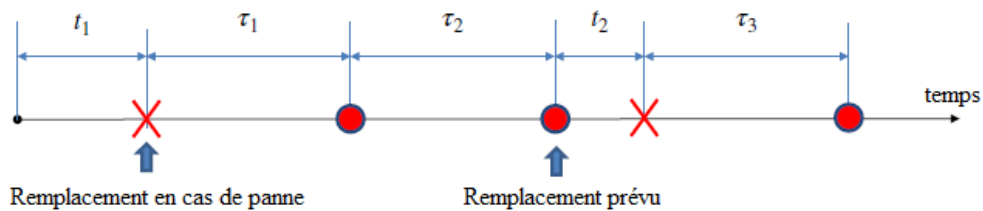


Figure I.3 Scénarios de remplacement pour un système réparable.

I.3.4 Durée de vie résiduelle moyenne

Dans l'ingénierie de la fiabilité, la durée de vie supplémentaire attendue, étant donné qu'un composant ou un système a survécu jusqu'au temps t , est appelée durée de vie résiduelle moyenne [5]. Si T la durée de vie d'un composant ou d'un système, la durée de vie résiduelle moyenne, dénommée $L(t)$, est donnée par :

$$L(t) = E[T - t | T \geq t] = \int_0^{\infty} (x - t) f_{T|T \geq t}(x) dx \quad (I.18)$$

Où $f_{T|T \geq t}(x)$ est le PDF conditionnel étant donné que $T \geq t$. La valeur de $L(t)$ peut être prédite ou estimée sur la base des données historiques de défaillance, et le résultat est fréquemment utilisé pour l'approvisionnement en pièces détachées ou l'allocation des ressources de réparation dans l'atelier de réparation. Le PDF conditionnel $f_{T|T \geq t}(x)$ peut être exprimé comme suit :

$$f_{T|T \geq t}(x) = \frac{f(x)}{R(t)} \quad (I.19)$$

En substituant l'équation (I.19) à (I.18), la durée de vie résiduelle moyenne est obtenue comme suit :

$$L(t) = \int_t^{\infty} (x - t) \frac{f(x)}{R(t)} dx = \frac{1}{R(t)} \int_t^{\infty} R(x) dx \quad (I.20)$$

I.4 Temps d'arrêt et disponibilité du système

I.4.1 Temps moyen de réparation (Mean-Time-to-Repair)

Le temps moyen de réparation (MTTR) représente le temps qui s'écoule entre le moment où le système est hors service et celui où il est utilisé. Le MTTR englobe le temps d'attente pour les techniciens de réparation, le délai de réception des pièces de rechange, le temps de diagnostic des défaillances, le temps de remplacement des pièces défectueuses et les autres temps d'arrêt liés aux inspections, ou des retards administratifs. Le MTTR est donnée par :

$$MTTR = t_{ad} + t_{pt} + t_{tn} + t_{ho} + t_{ft} \quad (I.21)$$

t_{ad} = retard administratif

t_{pt} = délai de réception de la pièce de rechange

t_{tn} = temps de montage de l'équipe de techniciens de réparation

t_{ho} = temps pratique pour le remplacement des unités défaillantes

t_{ft} = diagnostic des défaillances et durée des tests

I.4.2 Disponibilité du système

La disponibilité est la proportion du temps durant lequel un système est en état de fonctionnement. La fiabilité et la maintenabilité déterminent conjointement la disponibilité du système. En particulier, la première détermine la durée du MTBF et la seconde influence le MTTR. Elles sont liées à la disponibilité du système par la formule suivante :

$$A = \frac{MTBF}{MTBF + MTTR} \quad (I.22)$$

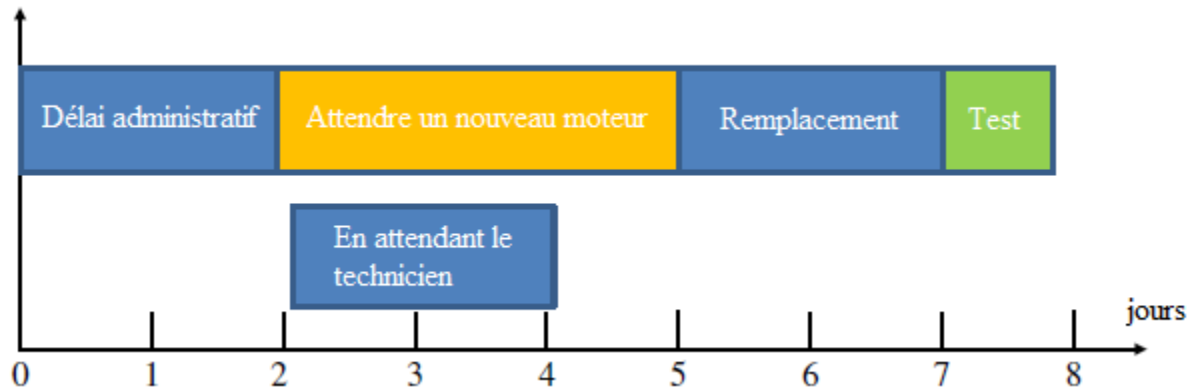


Figure I.4 Le MTTR du remplacement d'un moteur d'avion.

Il convient de mentionner que deux systèmes peuvent avoir la même disponibilité, mais que leur MTBF et MTTR peuvent être différents.

I.5 Conclusion

Comme nous l'avons précisé précédemment, l'analyse de la fiabilité est un outil très important pour caractériser le comportement du produit dans les différentes phases de vie, mesurer l'impact des modifications de conception sur l'intégrité du produit, qualifier un nouveau produit et améliorer ses performances tout au long de sa mission. Pour ça, nous avons montré la relation entre la fiabilité et les mesures probabilistes, Ensuite nous avons précisé quelques notions de durée de vie (MTTF, MTBF, MTBR, MTTR et Durée de vie résiduelle moyenne).

Chapitre II : Lois de fiabilité et tests d'ajustement

II.1 Introduction

L'analyse de la fiabilité est un outil très important pour caractériser le comportement du produit dans les différentes phases de vie, mesurer l'impact des modifications de conception sur l'intégrité du produit, qualifier un nouveau produit et améliorer ses performances tout au long de sa mission.

Dans les études de fiabilité, de nombreuses lois statistiques sont utilisées pour analyser les données de durée de vie. Ils sont suivis par des tests statistiques, ces tests sont des démarches conduisant à élaborer des règles de décision permettant de faire un choix entre deux hypothèses.

II.2 Méthodes paramétrique

II.2.1 Loi exponentielle

La variable aléatoire de la loi exponentielle est continue et non négative. C'est donc une variable aléatoire idéale pour modéliser la durée de vie des produits et du système. Si T est distribué de manière exponentielle, alors la PDF est donnée comme suit :

$$f(t) = \lambda e^{-\lambda t}, \text{ pour } t \geq 0 \quad (\text{II.1})$$

Où λ ($\lambda > 0$) est le paramètre de distribution. La CDF (fonction cumulée de densité ou fonction de répartition) et la fonction de fiabilité sont définies comme :

$$F(t) = P\{T \leq t\} = \int_0^t \lambda e^{-\lambda x} dx = 1 - e^{-\lambda t} \quad (\text{II.2})$$

$$R(t) = 1 - F(t) = e^{-\lambda t} \quad (\text{II.3})$$

La fonction du taux de défaillance et la fonction du taux de défaillance cumulée sont toutes deux définies comme suit :

$$h(t) = \frac{f(t)}{R(t)} = \lambda \quad (\text{II.4})$$

$$H(t) = \int_0^t \lambda dx = \lambda t \quad (\text{II.5})$$

Pour une distribution exponentielle de la durée de vie, la fonction du taux de défaillance est une constante, Ce qui lui confère une propriété dite 'sans mémoire'. Soit la durée t (en heures, par exemple) pendant laquelle le système a survécu, pour une variable aléatoire T possédant la propriété sans mémoire, l'égalité suivante est toujours valable :

$$P\{T > t + s | T > s\} = P\{T > t\}, \quad \text{pour } t > 0 \quad (\text{II.6})$$

Cette relation est issue de :

$$P\{T > t + s | T > s\} = \frac{P\{T > t + s | T > s\}}{P\{T > s\}} = \frac{P\{T > t + s\}}{P\{T > s\}} = \frac{e^{-\lambda(t+s)}}{e^{-\lambda s}} = e^{-\lambda t} = P\{T > t\} \quad (\text{II. 7})$$

Cela indique que la probabilité que le produit survive pendant $s + t$ heures, étant donné qu'il a survécu s heures, est la même que la probabilité initiale qu'il survive pendant t heures. Enfin, la moyenne et la variance de T sont obtenues et données comme suit :

$$E(T) = \int_0^{\infty} t f(t) dt = \int_0^{\infty} t \lambda e^{-\lambda t} dt = \frac{1}{\lambda} \quad (\text{II. 8})$$

$$E(T^2) = \int_0^{\infty} t^2 f(t) dt = \int_0^{\infty} t^2 \lambda e^{-\lambda t} dt = \frac{2}{\lambda^2} \quad (\text{II. 9})$$

$$V(T) = E(T^2) - E(T)^2 = \frac{1}{\lambda^2} \quad (\text{II. 10})$$

II.2.2 Loi de Weibull

L'idée de rechercher une distribution de la variable durée de vie X qui traduit les formes de trois parties de la courbe en baignoire est naturelle. W. Weibull a ainsi proposé dès 1958 de représenter la fonction taux de défaillance par une fonction en « t puissance β » décroissante pour $0 \leq \beta \leq 1$, constante pour $\beta = 1$ et croissante pour $\beta > 1$. Cette distribution, mathématiquement « praticable », s'est avérée une représentation satisfaisante du phénomène « temps de fonctionnement jusqu'à défaillance » dans de nombreuses situations concrètes.

On dit que la variable aléatoire X « durée de vie » obéit à une distribution de Weibull de paramètre de forme β , de paramètre de position γ et de paramètre d'échelle η si :

- elle est définie pour $t \geq \gamma$

- sa fonction taux de défaillance instantané λ **figure II.1** vérifie :

$$h(t) = \lambda(t) = \frac{\beta}{\eta} \left(\frac{t - \gamma}{\eta} \right)^{\beta - 1} \quad (\text{II. 11})$$

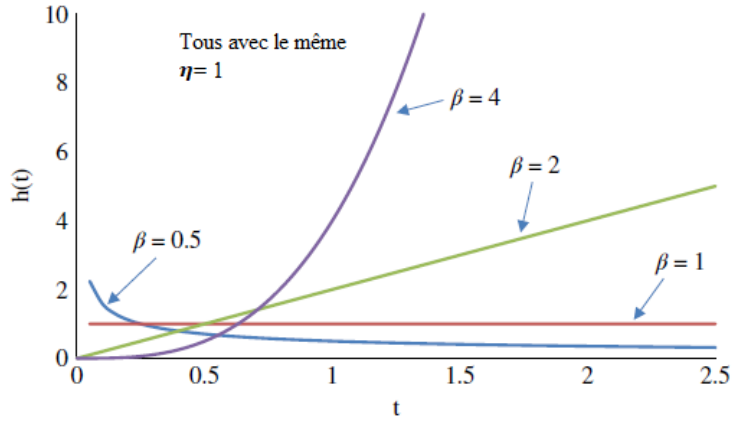


Figure II.1 Le taux de défaillance de la fonction de Weibull.

La fonction de densité $f(t)$, la fonction de fiabilité $R(t)$, **Figure II.2** et la MTBF sont données par :

$$\begin{cases} f(t) = \frac{\beta}{\eta} \left(\frac{t-\gamma}{\eta}\right)^{\beta-1} \cdot e^{-\left(\frac{t-\gamma}{\eta}\right)^\beta} \\ R(t) = e^{-\left(\frac{t-\gamma}{\eta}\right)^\beta} = 1 - F(t) \\ MTBF = \gamma + \eta\Gamma\left(1 + \frac{1}{\beta}\right) \end{cases} \quad (\text{II. 12})$$

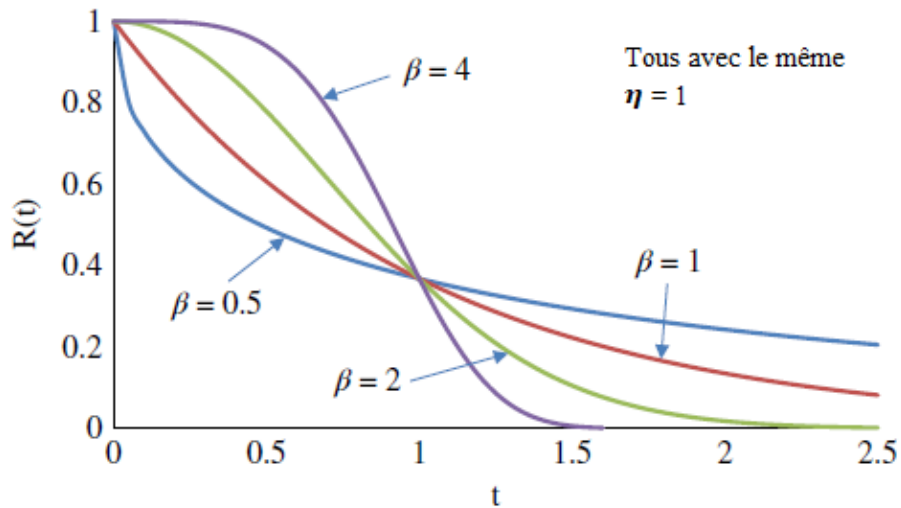


Figure II.2 Fonction de fiabilité de Weibull.

où $\Gamma(x)$ est la fonction intégrale Eulérienne de seconde espèce définie par la relation :

$$\Gamma(x) = \int_0^x (t^{x-1} \cdot e^{-t}) dt \quad (\text{II. 13})$$

L'algorithme de Pike et Hill (1966) permet un calcul extrêmement rapide de $\Gamma(x)$ à 10 décimales à partir d'un développement généralisé de $\ln(\Gamma(x))$. Cet algorithme est maintenant disponible comme "fonction – assistant " dans les tableurs.

La variance et le mode de cette distribution sont définies par :

$$V(X) = \eta^2 \left[\Gamma\left(1 + \frac{2}{\beta}\right) - \left(\Gamma\left(1 + \frac{1}{\beta}\right)\right)^2 \right] \quad (\text{II. 14})$$

$$\text{mode} = 0 \text{ si } \beta < 1 \text{ mais pour } \beta \geq 1 \text{ mode} = \eta \left(\frac{\beta - 1}{\beta}\right)^{1/\beta} \quad (\text{II. 15})$$

La valeur des paramètres β , γ et η affecte considérablement l'allure de la loi Weibull, ainsi :

Le paramètre de forme β (réel positif sans dimension) traduit l'allure du processus de dégradation du matériel :

- Si $0 \leq \beta \leq 1$ la fonction du taux de défaillance instantané est décroissante comme dans la partie « défaillance de jeunesse ou déverminage » de la courbe en baignoire.
- Si $\beta = 1$ on retrouve la distribution exponentielle.
- Si $\beta \geq 1$ le taux de défaillance instantané est une fonction croissante, on a affaire à un phénomène d'usure donc d'accélération du risque de défaillance avec le temps.

Le paramètre de position γ (unité : temps ou autre unité d'usage) est :

- Nul si les défaillances peuvent débuter à l'instant 0 (ou au nombre 0 d'unité d'usage).
- Positif si les défaillances ne peuvent se produire avant l'instant γ c'est-à-dire s'il y a nécessairement survie entre l'instant 0 et l'instant γ .
- Négatif si les défaillances ont débuté avant l'origine des temps choisie pour la mesure.

Le paramètre η (réel positif, unité : temps ou autre unité d'usage) n'est qu'un paramètre d'échelle, fonction de l'unité choisie.

Toutefois, il est toujours possible, par le changement de variable $t' = (t - \gamma)/\eta$, de travailler avec une distribution de Weibull réduite du type ($\beta > 0, \gamma = 0, \eta = 1$). Ce choix, nous l'avons fait dans les représentations suivantes de la fonction densité **Figure II.3** et de la fonction taux de défaillance instantané **Figure II.4**.

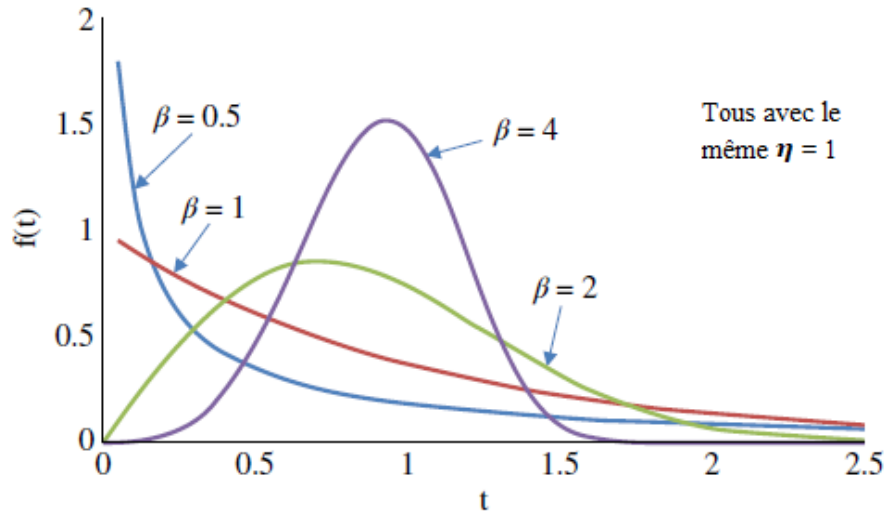


Figure II.3 Fonction de densité de probabilité de Weibull.

On constate que, pour β voisin de 3,5, la fonction densité est sensiblement symétrique, donc proche de la loi normale. En fait, la médiane de la distribution de Weibull est égale à la moyenne pour $\beta \cong 3,439\ 38$. Il existe de nombreux résultats relatifs à l'approximation de la loi de Weibull par la loi normale. A. Plait, par exemple, a montré que la fonction densité de Weibull ($\beta = 3.25, \gamma = 0, \eta = 1$) est proche de celle de la loi Normale ($\mu = 0,8693 ; \sigma = 0.303$) mais des conditions d'approximations bien plus fines ont été obtenues par Makino [6].

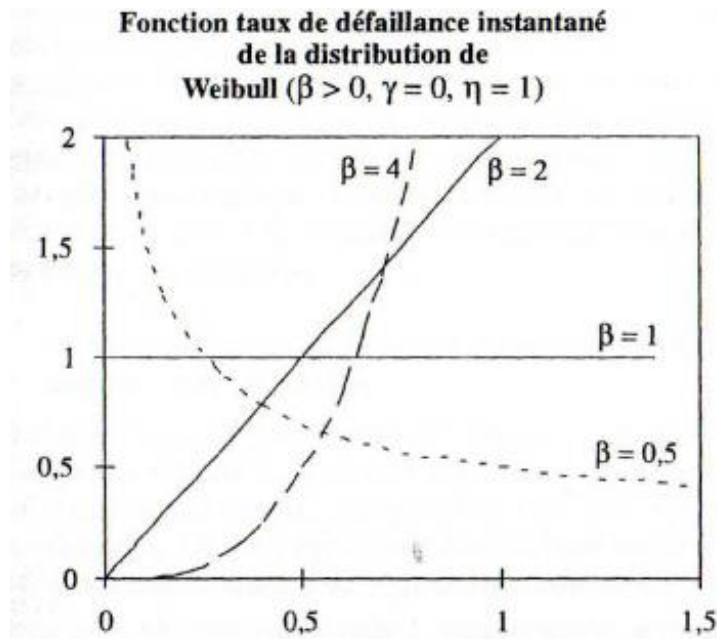


Figure II.4 Fonction de taux de défaillance instantané de la distribution de Weibull.

II.2.3 Loi normale

La loi normale, également connue sous le nom de loi gaussienne, est peut-être le modèle de distribution continue le plus utilisé dans les domaines de l'ingénierie et des sciences, y compris en analyse de fiabilité. On dit qu'une variable aléatoire T suit une distribution normale si sa PDF est décrite par :

$$f(t) = \frac{1}{\sqrt{2\pi\sigma}} e^{-\left(\frac{(t-\mu)^2}{2\sigma^2}\right)}, -\infty < t < \infty \quad (\text{II. 16})$$

Où μ et σ sont respectivement le paramètre de l'échelle et de forme. La fonction de densité normale a une courbe en forme de cloche qui est symétrique autour de μ . **Figure II.5** représente la PDF normale pour $\{\mu = 20, \sigma = 3\}$, $\{\mu = 20, \sigma = 5\}$ et $\{\mu = 35, \sigma = 5\}$ à des fins de comparaison.

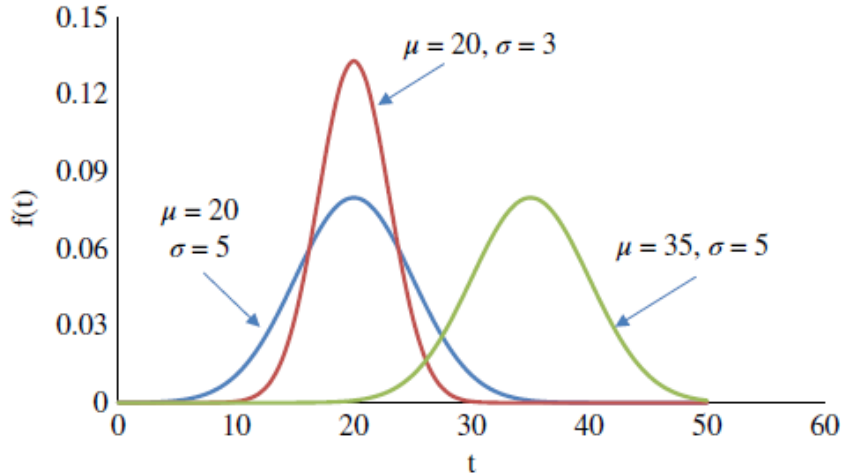


Figure II.5 la PDF normale pour différentes moyennes et variances.

La moyenne et la variance de T sont égales aux paramètres μ et σ^2 , respectivement,

$$E[T] = \int_{-\infty}^{\infty} tf(t)dt = \int_{-\infty}^{\infty} t \frac{1}{\sqrt{2\pi\sigma}} e^{-\left(\frac{(t-\mu)^2}{2\sigma^2}\right)} dt = \mu \quad (\text{II. 17})$$

$$V(T) = E[T^2] - E[T]^2 = \sigma^2 \quad (\text{II. 18})$$

La fonction de fiabilité est

$$R(t) = \int_t^{\infty} f(x)dx = 1 - \int_{-\infty}^t \frac{1}{\sqrt{2\pi\sigma}} e^{-\left(\frac{(x-\mu)^2}{2\sigma^2}\right)} dx = 1 - \Phi\left(\frac{t-\mu}{\sigma}\right) \quad (\text{II. 19})$$

Où $\Phi(z)$ désigne la fonction de répartition (la distribution cumulative normale standard) avec $\mu = 0$ et $\sigma = 1$. Contrairement à la distribution de Weibull, il n'y a pas d'expression de forme analytique pour l'équation (II. 19). Pour estimer $R(t)$ à un t donné, des tableaux sont créés pour énumérer la valeur cumulative possible pour $\Phi(z)$.

II.2.4 Loi log-normale

La distribution log-normale est l'une des distributions paramétriques les plus fréquemment utilisées dans l'analyse des données de fiabilité des dispositifs microélectroniques dans l'industrie de fabrication des semi-conducteurs. La raison pour laquelle les données sur la durée de vie des semi-conducteurs correspondent bien à la distribution log-normale est que celle-ci est formée par les effets multiplicatifs de variables aléatoires. Ce type d'interactions multiplicatives est souvent rencontré dans de nombreux mécanismes de défaillance des semi-conducteurs [7] [8]. On dit qu'une variable aléatoire T suit une distribution log-normale lorsque la variable transformée $X = \log(T)$ est normalement distribuée. La PDF pour T est donnée par :

$$f(t) = \frac{1}{t\sqrt{2\pi}\sigma} e^{-\left(\frac{(\log(t)-\mu)^2}{2\sigma^2}\right)}, \quad \text{pour } t \geq 0 \quad (\text{II. 20})$$

Deux paramètres μ et σ sont nécessaires pour définir une distribution log-normale, où μ est appelé le paramètre d'échelle et σ est le paramètre de forme. Contrairement à la distribution normale où la moyenne et l'écart-type, respectivement, sont égaux au paramètre de l'échelle et de la forme, la variance de la variable aléatoire log-normale est définie par :

$$E[T] = \int_0^{\infty} t \frac{1}{t\sqrt{2\pi}\sigma} e^{-\left(\frac{(\log(t)-\mu)^2}{2\sigma^2}\right)} dt = e^{\mu + \frac{\sigma^2}{2}} \quad (\text{II. 21})$$

$$E[T^2] = \int_0^{\infty} t^2 \frac{1}{t\sqrt{2\pi}\sigma} e^{-\left(\frac{(\log(t)-\mu)^2}{2\sigma^2}\right)} dt = e^{2\mu + 2\sigma^2} \quad (\text{II. 22})$$

$$V[T] = E[T^2] - E[T]^2 = e^{2\mu + \sigma^2} \cdot (e^{\sigma^2} - 1) \quad (\text{II. 23})$$

La **Figure II.6**. Représente la distribution log-normale pour $\sigma = 0,25, 0,5, \text{ et } 1$ avec $\mu = 2$. La PDF log-normal avec un σ plus petit a tendance à ressembler à la forme en cloche d'une loi normale.

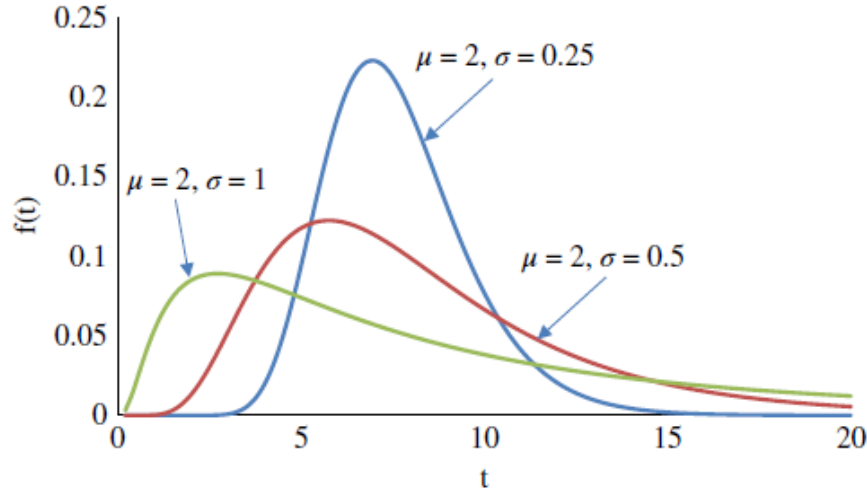


Figure II.6 Représentation de la FDP de loi Log-normal.

II.2.5 Loi Gamma

La distribution gamma est souvent utilisée pour modéliser le nombre d'erreurs dans les modèles de régression de Poisson à plusieurs niveaux, car la combinaison de la distribution de Poisson et d'une distribution gamma est une distribution binomiale négative. Dans les statistiques de fiabilité bayésiennes, la distribution gamma est souvent choisie comme un a priori conjugué pour le paramètre de distribution à estimer. Par exemple, la distribution gamma est le conjugué a priori pour la distribution exponentielle de la durée de vie. La PDF d'une variable aléatoire T avec la distribution gamma est donné comme suit

$$f(t) = \frac{\lambda e^{-\lambda t} \cdot (\lambda t)^{\theta-1}}{\Gamma(\theta)}, \text{ pour } t \geq 0 \quad (\text{II. 24})$$

Avec

$$\Gamma(\theta) = \int_0^{\infty} e^{-x} x^{\theta-1} dx, \text{ pour } \theta > 0 \quad (\text{II. 25})$$

Où λ et θ sont des valeurs positives. L'équation (II. 25) est appelée la fonction gamma. Si θ est un nombre entier ou négatif, alors :

$$\Gamma(\theta) = (\theta - 1)!$$

La moyenne et la variance de la variable aléatoire gamma est [9]:

$$E[T] = \frac{\theta}{\lambda} \quad (\text{II. 26})$$

$$V[T] = \frac{\theta}{\lambda^2} \quad (\text{II.27})$$

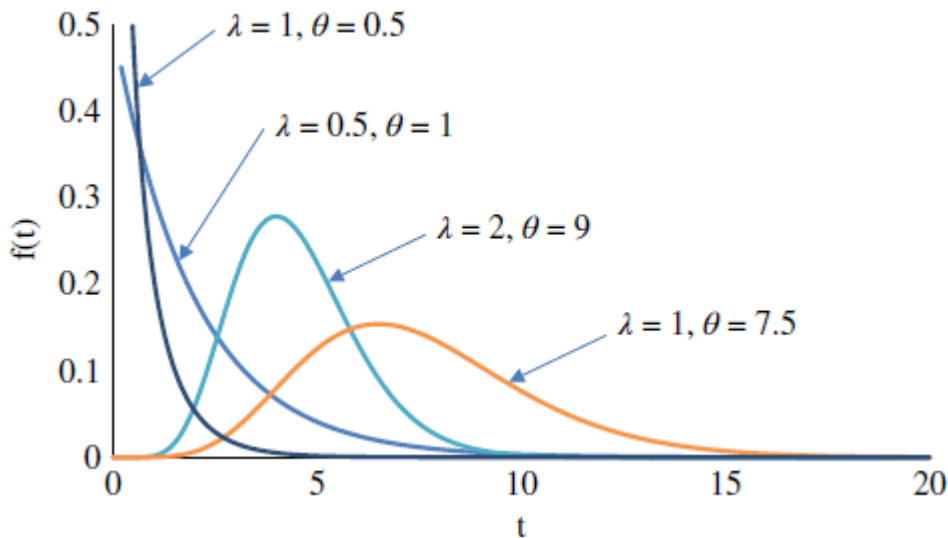


Figure II.7 Représentation de la PDF gamma pour différentes valeurs de λ et θ .

La **Figure II.7** représente la PDF gamma avec différentes paires de λ et θ . Pour $0 < \theta \leq 1$, la fonction PDF diminue de façon monotone avec le temps t . Lorsque $\theta > 1$, la PDF a plutôt une forme en cloche. En faisant varier les valeurs de λ et θ , on peut obtenir différentes formes de la courbe PDF gamma. Tout comme la loi de Weibull, la loi gamma est également capable de modéliser un taux de défaillance décroissant, constant et croissant ce qui correspond respectivement à $0 < \theta < 1$, $\theta = 1$, et $\theta > 1$.

Pour le cas particulier où les valeurs de θ sont des entiers positifs, la loi gamma représente la distribution du temps qu'il faudra attendre pour qu'un total de n événements se produisent. Si T_n indique le moment où le $n^{\text{ème}}$ événement se produit, alors T_n est inférieur ou égal à t si et seulement si le nombre d'événements survenus avant t est au moins n , à savoir :

$$P\{T_n \leq t\} = P\{N(t) \geq n\} = \sum_{k=n}^{\infty} P\{N(t) = k\} = \sum_{k=n}^{\infty} \frac{e^{-\lambda t} (\lambda t)^k}{k!} \quad (\text{II.28})$$

En prenant la dérivée de cette relation par rapport à t , on obtient la PDF de T_n comme suit :

$$f(t) = \frac{\lambda (\lambda t)^{n-1} \cdot e^{-\lambda t}}{(n-1)!} \quad (\text{II.29})$$

Ainsi, T_n est une loi gamma avec les paramètres (λ, n) . Cette distribution est souvent appelée loi n-Erlang, qui est un cas particulier de loi gamma avec θ étant des entiers positifs. Si $\theta = 1$, la distribution gamma est réduite à une distribution exponentielle.

II.2.6 Loi d'extremum généralisée

La théorie des valeurs extrêmes a pour objectif essentiel la maîtrise des risques. Elle traite le comportement stochastique des valeurs extrêmes dans un processus [10].

Supposons (X_1, \dots, X_n) des variables aléatoires indépendantes et identiquement distribuées (i.i.d) avec fonction de répartition commune $F(x) = P[X_i \leq x]$. Soit $M_n = \max(X_1, \dots, X_n)$ le maximum des n premières variables aléatoires de notre échantillon. La fonction de répartition de la variable aléatoire M_n est donnée par la relation suivante :

$$F_{M_n}(x) = P[M_n \leq x] = F(x)^n \quad (\text{II. 30})$$

et donc $F_{M_n}(x)$ converge vers 0 ou 1 selon la valeur de $F(x) \in [0; 1]$. La loi limite de M_n est donc dégénérée et on autorise une renormalisation linéaire de M_n pour contourner le problème. Plus précisément, on s'intéresse au comportement limite de :

$$M_n^* = \frac{M_n - b_n}{a_n} \quad (\text{II. 31})$$

avec a_n et b_n des suites (déterministes) bien choisies pour éviter la dégénérescence de la loi limite. Il existe une analogie avec le théorème central-limite qui s'intéresse au comportement de $\frac{S_n - b_n}{a_n}$ avec S_n définie par $S_n = X_1 + \dots + X_n$, $b_n = n\mathbb{E}[X_1]$ et $a_n = n\text{Var}(X_1)$ sous certaines conditions, la loi limite de $\frac{S_n - b_n}{a_n}$ est une loi $\mathcal{N}(0,1)$

Le théorème de Fisher-Tippett montre que la seule loi limite de M_n^* est la loi GEV définie ci-dessous.

Distribution GEV : $X \sim \text{GEV}(\mu, \sigma, k)$ avec $\sigma > 0$ si sa fonction de répartition est donnée par

- Si $k \neq 0$: $F(\mu, \sigma, k) = \exp\left(-\left[1 + k \frac{x-\mu}{\sigma}\right]^{-\frac{1}{k}}\right)$ définie pour x vérifiant $1 + k \frac{x-\mu}{\sigma} > 0$
- Si $k = 0$: $F(\mu, \sigma, k) = \exp\left(-\exp\left(-\frac{x-\mu}{\sigma}\right)\right)$ définie pour tout $x \in \mathcal{R}$

Remarques :

- μ est un paramètre de position et σ paramètre d'échelle si : $X \sim \text{GEV}(\mu, \sigma, k)$ alors $\frac{x-\mu}{\sigma} \sim \text{GEV}(0,1, k)$.
- k est un paramètre de forme.
 - Le cas $k = 0$ correspond à la loi de Gumbel et on peut montrer que c'est bien la limite des autres cas lorsque $k \rightarrow 0$.

- Le cas $k < 0$ correspond à la loi de Weibull. La formule ci-dessus définit la fonction de répartition pour $x < (\mu - \sigma)/k = x_+$ et on pose $F(x, \mu, \sigma, k) = 1$ si $x \geq x_+$: la variable X est donc à valeurs dans $] - \infty, x]$ (loi à support majoré).
 - Le cas $k > 0$ correspond à la loi de Fréchet. La formule ci-dessus définit la fonction de répartition pour $x > (\mu - \sigma)/k = x_-$ et on pose $F(x, \mu, \sigma, k) = 0$ si $x \leq x_-$: la variable X est donc à valeurs dans $]x_-; \infty[$.
- la densité $f(x, \mu, \sigma, k)$ de la loi $GEV(\mu, \sigma, k)$ s'obtient en dérivant la fonction de répartition par rapport à x . Si $k \neq 0$, on obtient

$$f(x, \mu, \sigma, k) = \begin{cases} \frac{1}{\sigma} \left(1 + k \frac{x-\mu}{\sigma}\right)^{-\frac{1}{k-1}} \exp\left(-\left(1 + k \frac{x-\mu}{\sigma}\right)^{-\frac{1}{k}}\right) & \text{si } 1 + k \frac{x-\mu}{\sigma} > 0 \\ 0 & \text{sinon} \end{cases} \quad (\text{II. 32})$$

on en déduit que X a un moment d'ordre $k > 0$ si $k < \frac{1}{k}$. En particulier X a une espérance finie si $k < 1$ et une variance finie si $k < \frac{1}{2}$.

Théorème (Fisher-Tippet (1928))

S'il existe des séquences $(a_n)_{n \in \mathbb{N}}$ et $(b_n)_{n \in \mathbb{N}}$ telles que

$$P \left[\frac{M_n - b_n}{a_n} \leq z \right] \rightarrow G(z) \quad (\text{II. 33})$$

pour tout $z \in \mathbb{R}$ avec G non dégénérée, alors G est la fonction de répartition d'une loi GEV.

II.3 Méthodes non paramétrique

Il existe plusieurs méthodes d'estimation d'une fonction de densité de probabilité. Parmi ces méthodes non paramétriques, citons la méthode des histogrammes, celle des plus proches voisins, celle fondée sur des bases de fonctions, etc. Mais là plus utilisée est la méthode par noyaux dont nous présenterons quelques aspects. Elle a été introduite par Rosenblatt et développée par Parzen.

II.3.1 Estimation d'une fonction de densité de probabilité par noyaux

Soit X une v. a. qui admet une densité $f_X(t) = f(t)$. Nous nous proposons d'estimer cette dernière à partir de n échantillon de X .

Nous appelons noyau de Parzen toute fonction $k_e(t)$ positive, paire, bornée, satisfaisant aux conditions :

$$\lim_{t \rightarrow \infty} t k_e(t) = 0 \quad \text{et} \quad \int_{-\infty}^{\infty} k_e(t) dt < \infty \quad (\text{II. 34})$$

- **Remarque :**

Les conditions supplémentaires suivantes nous permettront d'obtenir des propriétés usuelles pour les estimateurs de densités :

- Le noyau est une densité : $\int_{-\infty}^{\infty} k_e(t) dt = 1$.
- Le noyau est de carré intégrable : $d^2(k_e) = \int_{-\infty}^{\infty} (k_e(t))^2 dt < \infty$.
- Le noyau admet un moment d'ordre deux : $\sigma^2(k_e) = \int_{-\infty}^{\infty} t^2 k_e(t) dt < \infty$.

Les noyaux les plus utilisés sont :

- Le noyau uniforme : $k_e(t) = \frac{1}{2} I_{[-1;1]}(t)$. Il correspond à la densité de la loi Uniforme $\mathcal{U}(-1; 1)$.
- Le noyau triangulaire : $k_e(t) = (1 - |t|) I_{[-1;1]}(t)$. Il correspond à la densité de la loi Triangulaire $\mathcal{TR}(0)$.
- Le noyau d'Epanechnikov : $k_e(t) = \frac{3}{4} (1 - t^2) I_{[-1;1]}(t)$.
- Le noyau quartique $k_e(t) = \frac{15}{16} (1 - t^2)^2 I_{[-1;1]}(t)$.
- Le noyau circulaire : $k_e(t) = \frac{\pi}{4} \cos\left(\frac{\pi}{2} t\right) I_{[-1;1]}(t)$.
- Le noyau gaussien : $k_e(t) = \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}}$. Il correspond à la densité de la loi Normale standard $\mathcal{N}(0; 1)$.

Toutes ces fonctions sont bien des noyaux et satisfont à l'ensemble des conditions de la Remarque.

Nous appelons estimateur de Rosenblatt-Parzen ou estimateur à noyau la v. a.

$$f_n^*(t) = \frac{1}{nh} \sum_{j=1}^n k_e\left(\frac{t - X_j}{h}\right) = \frac{1}{h} \mathbb{E} \left[k_e\left(\frac{t - X_{EM}}{h}\right) \right]. \quad (\text{II. 35})$$

où est un noyau de Parzen et le nombre $h = h_n > 0$ est un paramètre de lissage appelé fenêtre. Nous notons également $f_n^*(t)$ une réalisation de l'estimateur [11].

II.3.2 Méthode d'estimation de Kaplan-Meier

L'estimateur de Kaplan-Meier, également connu sous le nom de l'estimateur produit-limite, est un estimateur pour estimer la fonction de survie d'après des données de durée de vie. En recherche

médicale, il est souvent utilisé pour mesurer la fraction de patients en vie pour une certaine durée après leur traitement. Il est également utilisé en économie et en écologie.

Une courbe d'estimation de Kaplan-Meier pour la fonction de survie est une série de marches horizontales de grandeur décroissante qui, quand un échantillon suffisamment grand est utilisé, permet d'approcher la fonction de survie réelle dans cette population. La valeur de la fonction de survie entre les échantillons successifs observés est considérée comme étant constante. [12].

L'estimateur de Kaplan-Meier est considéré comme l'estimateur non paramétrique le plus direct de la fonction de survie. Il est l'unique estimateur cohérent de la fonction de survie.

Formules

Soit $S(t)$ la probabilité qu'un membre d'une population donnée ait une durée de vie supérieure à t . Pour un échantillon de taille n dans cette population, les durées observées jusqu'à chaque décès des membres de l'échantillon n sont :

$$t_1 \leq t_2 \leq t_3 \leq \dots \leq t_n$$

À chaque n_i correspond un t_i , n_i étant le nombre de personnes « à risque » juste avant le temps t_i , et d_i le nombre de décès au temps t_i .

L'estimateur de Kaplan-Meier est l'estimation du maximum de vraisemblance non-paramétrique de $S(t)$. C'est un produit de la forme :

$$\hat{S}(t) = \prod_{t_i < t} \frac{n_i - d_i}{n_i} \quad (\text{II. 38})$$

Lorsqu'il n'y a aucune censure, n_i est le nombre de survivants juste avant le temps t_i .

Lorsqu'il y a censure, n_i est le nombre de survivants moins le nombre de pertes (cas censurés). Ce sont seulement ces cas survivants qui continuent à être observés (qui n'ont pas encore été censurés) qui sont « à risque » de décès (observé).

Ici une autre définition possible parfois utilisée :

$$\hat{S}(t) = \prod_{t_i \leq t} \frac{n_i - d_i}{n_i} \quad (\text{II. 39})$$

Les deux définitions diffèrent uniquement aux moments des événements observés. La dernière définition est « continue à droite » tandis que la première est « continue à gauche ». Soit T la variable aléatoire qui mesure le temps d'échec et soit $F(t)$ sa Fonction de répartition cumulative. On note que :

$$S(t) = P[T > t] = 1 - P[T \leq t] = 1 - F(t) \quad (\text{II. 40})$$

En conséquence, la définition continue à gauche de $\hat{S}(t)$ peut-être préférée pour rendre l'estimation compatible avec une estimation continue à droite de $F(t)$.

L'estimateur de Kaplan-Meier est une statistique, et certains estimateurs sont utilisés pour approcher sa variance. Un de ces estimateurs les plus courants est la formule de Greenwood :

$$\widehat{Var}(\hat{S}(t)) = S^{\wedge}(t)^2 \sum_{t_i < t} \frac{d_i}{n_i(n_i - d_i)} \quad (\text{II. 41})$$

Un avantage important de la courbe de Kaplan-Meier est que cette méthode peut prendre en compte certains types de données censurées, en particulier censurées par la droite, ce qui intervient lorsqu'un patient disparaît d'une étude, c'est-à-dire qu'on ne dispose plus de ses données avant que l'évènement attendu (par exemple le décès), soit observé. Sur le graphique, les petits traits verticaux indiquent ces censures. Si aucune troncature ou censure n'intervient, la courbe de Kaplan-Meier est équivalente à la fonction de survie.

II.4 Test d'hypothèse

Une hypothèse statistique est un énoncé (une affirmation) concernant les caractéristiques (valeurs des paramètres, forme de la distribution des observations) d'une population.

Un test d'hypothèse (ou test statistique) est une démarche qui a pour but de fournir une règle de décision permettant, sur la base de résultats d'échantillon, de faire un choix entre deux hypothèses statistiques.

II.4.1 Le Test du Khi-Deux (khi square χ^2)

On peut distinguer trois types de test du χ^2 :

- le test du χ^2 d'adéquation (H_0 : « le caractère X suit-il une loi particulière ? »),
- le test du χ^2 d'homogénéité (H_0 : « le caractère X suit-il la même loi dans deux populations données ? »).
- le test du χ^2 d'indépendance (H_0 : « le caractère X et Y sont-ils indépendants ? »).

Ces trois tests ont un principe commun: on répartit les observations dans k classes dont les effectifs sont notés $n_1 = N_1(\omega), \dots, n_k = N_k(\omega)$. L'hypothèse H_0 permet de calculer les effectifs théoriques notés $n_{i\acute{e}me=1}, \dots, n_{k\acute{e}me}$. On rejette H_0 si les effectifs observés sont trop différents des effectifs théoriques.

Pour cela on utilise la statistique de test

$$U = \sum_{i=1}^k \frac{(N_i - n_{i\acute{e}me})^2}{n_{i\acute{e}me}} \quad (\text{II. 42})$$

Cette statistique suit la loi $U \simeq \chi^2(k - 1 - m)$ où k est le nombre de classes et m est le nombre de paramètres estimés nécessaire au calcul des effectifs théoriques.

Il faut s'assurer que les effectifs théoriques sont plus grands que 5 et donc faire des regroupements de classes si besoin. A partir de là, on calcule la zone de rejet unilatérale $R_\alpha =]t_\alpha; \infty[$ au risque α en déterminant t_α dans la table de $\chi^2(k - 1 - m)$ par :

$$P[U > t_\alpha] = \alpha \quad (\text{II. 43})$$

La règle de décision est la suivante :

- Si $u = \sum_{i=1}^k \frac{(N_i - n_{i\text{éme}})^2}{n_{i\text{éme}}}$ appartient à R_α on rejette H_0 .
- Si $u = \sum_{i=1}^k \frac{(N_i - n_{i\text{éme}})^2}{n_{i\text{éme}}}$ n'appartient pas à R_α on accepte H_0 .

Remarque

- Contrairement aux autres tests, les tests du χ^2 n'exigent pas de formuler l'hypothèse alternative H_1 , qui correspond à la négation de H_0 .
- Les effectifs théoriques doivent être supérieurs à 5. Si ce n'est pas le cas, il faut reconstituer les classes.
- Dans la statistique $U \simeq \chi^2(k - 1 - m)$, on manipule des effectifs et non des pourcentages.

II.4.2 Test d'Anderson-Darling

Le test d'Anderson-Darling est un test de (goodness-fit) qui permet de contrôler l'hypothèse selon laquelle la loi d'une variable aléatoire observée dans un échantillon suit une certaine loi théorique. Il permet notamment de tester si la loi empirique obtenue correspond à une loi normale.

Considérons la variable aléatoire X , qui suit la loi normale avec une espérance μ et une variance σ^2 , et qui a une fonction de distribution $F_X(x, \theta)$ où θ est un paramètre (ou un ensemble de paramètres) qui détermine, F_X . Nous supposons en outre que θ est connu.

Une observation d'un échantillon de taille n issu de la variable X donne une fonction de distribution $F_n(x)$. La statistique d'Anderson-Darling, désignée par A^2 , est alors donnée par la somme pondérée des carrés des écarts $F_X(x, \theta) - F_n(x)$:

$$A^2(x) = \frac{1}{n} \left(\sum_{i=1}^n (F_X(x, \theta) - F_n(x))^2 \right) \quad (\text{II. 44})$$

Partant du fait que A^2 est une variable aléatoire qui suit une certaine loi sur l'intervalle $[0 + \infty[$, il est possible de tester, pour un niveau de signification fixé a priori, si $F_n(x)$ est la réalisation de la variable aléatoire $F_X(x, \theta)$, c'est-à-dire si X suit la loi de probabilité avec la fonction de distribution $F_X(x, \theta)$.

Calcul de la statistique A^2

On dispose les observations x_1, x_2, \dots, x_n dans l'échantillon issu de X par ordre croissant, $x_1 < x_2 < \dots < x_n$. Notez que $z_i = F_X(x, \theta), (i = 1, 2, \dots, n)$. Ensuite, on calcule A^2 :

$$A^2 = -\frac{1}{n} \left(\sum_{i=1}^n (2i - 1) (\ln(z_i) + \ln(1 - z_{n+1-i})) \right) - n \quad (\text{II. 45})$$

Dans le cas où X suit une distribution normale avec moyenne μ et variance σ^2 , nous pouvons énumérer quatre cas, en fonction des paramètres connus μ et σ^2 (F est la fonction de distribution de la loi normale) [12] :

- 1- μ et σ^2 sont connus, donc $F_X(x; (\mu, \sigma^2))$ est parfaitement spécifiée. Naturellement, nous avons alors $z_i = F(w_i)$ où $w_i = \frac{x_i - \mu}{\sigma}$.
- 2- σ^2 est connu mais μ est inconnu et estimé en utilisant $\bar{x} = \frac{1}{n} \sum_i x_i$ la moyenne de l'échantillon. Ensuite, on laisse $z_i = F(w_i)$, où $w_i = \frac{x_i - \bar{x}}{s}$.
- 3- μ est connu mais σ^2 est inconnu et est estimé à l'aide de $S^2 = \frac{1}{n} (\sum_i (x_i - \mu)^2)$. Dans ce cas, laissez $z_i = F(w_i)$, où $w_i = \frac{x_i - \mu}{s}$.
- 4- μ et σ^2 sont tous deux inconnus et sont estimés respectivement à l'aide de \bar{x} et $S^2 = \frac{1}{n-1} (\sum_i (x_i - \bar{x})^2)$. Ensuite, on laisse $z_i = F(w_i)$, où $w_i = \frac{x_i - \bar{x}}{s}$.

Des distributions asymptotiques ont été trouvées pour A^2 par Anderson et Darling pour le premier cas, et par Stephens pour les deux cas suivants. Pour le dernier cas, Stephens a déterminé une distribution asymptotique pour la transformation :

$$A^* = A^2 \left(1,0 + \frac{0,75}{n} + \frac{0,75}{n^2} \right) \quad (\text{II. 48})$$

Par conséquent, comme indiqué ci-dessous, nous pouvons construire un tableau qui donne, selon le cas et le niveau de signification (10%, 5%, 2,5% ou 1%), les valeurs limites de A^2 (et A^* pour le quatrième cas) au-delà desquelles l'hypothèse de normalité est rejetée :

Cas :	Niveau de signification			
	0.1	0.050	0.025	0.01
1: $A^2 =$	1.933	2.492	3.070	3.857
2: $A^2 =$	0.894	1.087	1.285	1.551
3: $A^2 =$	1.743	2.308	2.898	3.702
4: $A^* =$	0.631	0.752	0.873	1.035

Figure II.8 tableau de niveau de signification.

Domaines et Limitations

Comme la distribution de A^2 est exprimée de manière asymptotique, le test a besoin que la taille de l'échantillon n soit importante. Si ce n'est pas le cas, alors, pour les deux premiers cas, la distribution de A^2 n'est pas connue et il est nécessaire d'effectuer une transformation du type $A^2 \rightarrow A^*$, à partir de laquelle A^* peut être déterminé. Lorsque $n > 20$, on peut éviter une telle transformation et les données du tableau ci-dessus sont donc valables.

Le test d'Anderson-Darling présente l'avantage de pouvoir être appliqué à un large éventail de lois (pas seulement une loi normale, mais aussi des lois exponentielles, logistiques et gamma, entre autres). Cela nous permet d'essayer un large éventail de distributions alternatives si le test initial rejette l'hypothèse nulle pour la distribution d'une variable aléatoire [13].

II.4.3 Test de Kolmogorov-Smirnov (1933, 1939)

On suppose que la fonction de répartition F de la variable aléatoire X est continue et strictement croissante. Soit F^* la fonction de répartition empirique d'un échantillon de taille n issu de cette population.

Variable de décision

La variable de décision est la variable aléatoire D_n définie par :

$$D_n = \text{Sup}_{x \in \mathbb{R}} |F^*(x) - F(x)| \tag{II.49}$$

Glivenko et Kolmogorov ont démontré que la fonction $K_n(y)$ définie par :

$$K_n(y) = \text{Pr}[\sqrt{n}D_n < y] \tag{II.40}$$

Converge, quand n tend vers l'infini, vers une fonction $K(y)$:

$$\begin{cases} K(y) = 0 & y \leq 0 \\ K(y) = \sum_{k=-\infty}^{\infty} (-1)^k e^{-2k^2 y^2} & y > 0 \end{cases} \tag{II.50}$$

Des tables donnent les valeurs de cette fonction K .

Règle de décision

On rejette l'hypothèse H_0 si la valeur de la statistique, obtenue à partir de l'échantillon, est supérieure à une valeur $d(n)$ n'ayant qu'une probabilité α d'être dépassée.

Sinon, on garde l'hypothèse H_0 et on considère que la distribution théorique spécifiée est acceptable, c'est-à-dire $F = F_0$.

Remarques

Le test de Kolmogorov-Smirnov est préférable au test du chi-deux pour des variables continues. En effet, la variable aléatoire de décision D_n utilise l'échantillon tel qu'il se présente, en revanche, le test du chi-deux appauvrit l'information en regroupant les données par classes et en assimilant les données d'une classe à la valeur centrale.

II.4.4 Test de Cramer-Von-Mises

On considère la statistique $n\omega^2$ définie par :

$$n\omega^2 = \int_{-\infty}^{\infty} [F^*(x) - F(x)] dF(x) \quad (\text{II. 51})$$

Il existe des tables pour la loi de cette variable aléatoire, loi indépendante de F . Elle est utilisée pour évaluer l'écart entre une distribution empirique et une distribution théorique. Si les valeurs x_i de l'échantillon sont ordonnées par valeurs croissantes, on démontre que :

$$n\omega^2 = \frac{1}{12n} \sum_{i=1}^n \left[\frac{2i-1}{2n} - F(x_i) \right]^2 \quad (\text{II. 52})$$

On rejette l'hypothèse H_0 si la valeur de la statistique $n\omega^2$ donnée par l'échantillon est supérieure à une valeur n'ayant qu'une probabilité α d'être dépassée. Pour $\alpha = 0,05$, on rejette l'hypothèse H_0 si $n\omega^2 > 0,46136$.

II.4.5 Tests de normalité et d'exponentialité

Si les paramètres des lois de distribution ne sont pas connus mais estimés, on peut utiliser les résultats empiriques suivants (Biometrika Tables) [14].

Test de normalité

- Hypothèse H_0 . Loi normale $\mathcal{N}(m; s)$

m est estimée par \bar{x} et l'écart-type s par : $\sqrt{\frac{1}{n-1} \sum_i (x_i - \bar{x})^2}$.

- Règle de décision. On rejette l'hypothèse H_0 :
au seuil $\alpha = 0,005$ si $\left(\sqrt{n} + \frac{0,85}{\sqrt{n}} - 0,01\right) D_n > 0,895$.
ou si $\left(1 + \frac{0,5}{n}\right) n\omega^2 > 0,126$.
au seuil $\alpha = 0,01$ si $\left(\sqrt{n} + \frac{0,85}{\sqrt{n}} - 0,01\right) D_n > 1,035$.
ou si $\left(1 + \frac{0,5}{n}\right) n\omega^2 > 0,178$.

Test d'exponentialité

- Hypothèse H_0 . Loi exponentielle de densité $f(x) = \frac{1}{\theta} \exp\left(-\frac{x}{\theta}\right)$.
 θ est estimée par \bar{x} .
- Règle de décision. On rejette l'hypothèse H_0 :
au seuil $\alpha = 0,05$ si $\left(\sqrt{n} + \frac{0,5}{\sqrt{n}} + 0,26\right) \left(D_n - \frac{0,2}{n}\right) > 1,094$.
ou si $\left(1 + \frac{1,6}{n}\right) n\omega^2 > 0,224$
au seuil $\alpha = 0,01$ si $\left(\sqrt{n} + \frac{0,5}{\sqrt{n}} + 0,26\right) \left(D_n - \frac{0,2}{n}\right) > 1,304$.
ou si $\left(1 + \frac{1,6}{n}\right) n\omega^2 > 0,337$.

II.5 Conclusion

Dans ce chapitre, nous avons mis l'accent sur les lois de probabilité principales soit paramétrique et non paramétrique plus utilisé dans l'analyse de fiabilité. Aussi on a précisé chacune loi leur principe. Enfin, nous avons terminé ce chapitre par présentation des quelques tests statistiques pour le but d'estimation de fiabilité.

Chapitre III : Présentation de la toolbox et application



III.1 Introduction

MATLAB est considéré par de nombreux utilisateurs non seulement comme un langage de haute performance pour l'informatique technique, mais aussi comme un environnement pratique pour la construction d'interfaces graphiques (GUI). La visualisation des données et la conception de l'interface utilisateur graphique dans MATLAB sont basées sur le système graphique Handle dans lequel les objets organisés dans une hiérarchie d'objets graphiques peuvent être manipulés par diverses commandes de haut et de bas niveau.

III.2 Toolbox en Matlab

'Statistics and Machine Learning Toolbox' fournit des fonctions et des applications pour décrire, analyser et modéliser les données. Elle permet d'utiliser des statistiques descriptives, des visualisations et des regroupements pour l'analyse exploratoire des données, ajuster des distributions de probabilité aux données, générer des nombres aléatoires pour les simulations de Monte Carlo et effectuer des tests d'hypothèse. Les algorithmes de régression et de classification permettent de tirer des conclusions à partir des données et de construire des modèles prédictifs, soit de manière interactive, en utilisant les applications de classification et de régression, soit de manière programmatique [15].

III.2.1 Les interfaces graphiques GUI (Graphical User Interface)

Les applications MATLAB sont des programmes autonomes dotés d'une interface graphique, qui permettent d'automatiser une tâche ou un calcul. L'interface graphique contient généralement des commandes telles que des menus, des barres d'outils, des boutons et des curseurs. De nombreux produits MATLAB, comme 'Curve Fitting Toolbox', 'Signal Processing Toolbox' et 'Control System Toolbox', incluent des applications dotées d'une interface graphique personnalisée. Il est également possible de créer ses propres applications personnalisées, ainsi que leur interface graphique, afin de permettre à d'autres de les utiliser.

Le développement des interfaces graphiques peut être séparé en deux parties :

- Gestion de la mise en place et des propriétés des objets.
- Programmation des interactions avec les objets.

Il existe deux méthodes de développement des interfaces graphiques sous MATLAB. La première utilise un outil graphique dédié, la seconde nécessite de programmer entièrement par lignes de code, ce qui est le cas de la toolbox développée de fiabilité [16].

III.2.2 L'outil GUIDE

Depuis la version 5.0 (1997), MATLAB possède un outil dédié à la création des interfaces graphiques appelé GUIDE (pour Graphical User Interface Development Environment). Le GUIDE est un constructeur d'interface graphique qui regroupe tous les outils dont le programmeur a besoin

pour créer une interface graphique de façon intuitive. Le placement des objets est réalisé par sélection dans une boîte à outils, et leur mise en place se fait à l'aide de la souris.

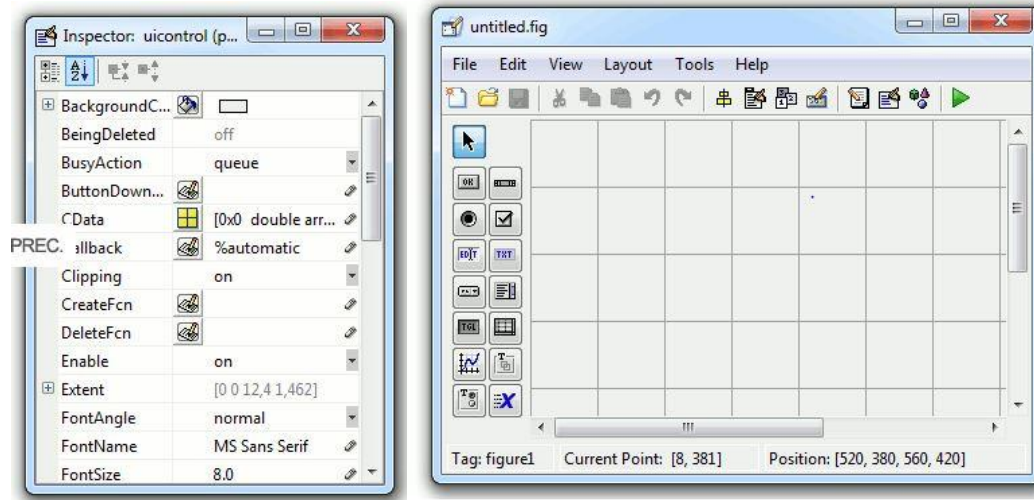


Figure III.1 L'outil GUIDE de Matlab.

Un double-clic sur un objet permet de faire apparaître le 'Property Inspector' où les propriétés des objets sont facilement éditables. Leurs modifications et la visualisation de ces modifications sont immédiates. Le GUIDE possède également des outils pour gérer l'alignement des objets et pour créer des barres d'outils ou des menus.

Le fichier 'fig' contient la définition des objets graphiques (positions et propriétés). Ce fichier peut être ouvert ultérieurement avec le GUIDE pour modifier les objets graphiques. Quant au fichier '.m', il contient les lignes de code qui assurent le fonctionnement de l'interface graphique (actions des objets). Ce fichier peut être édité dans 'MATLAB Editor' pour y ajouter des actions à la main. C'est ce fichier qui doit être lancé pour utiliser l'interface graphique [17].

III.3 Présentation de la toolbox de fiabilité

Ce programme de fiabilité est développé sous Matlab et peut être utilisé dans tout domaine impliquant l'analyse de fiabilité. À partir des temps de défaillance on pourra établir la fonction de survie, la fonction de densité de probabilité et la fonction hasard (ou risque). Un ajustement sur des lois de distribution statistique (exponentiel, normal, Weibull, ...) est réalisé en tenant compte des tests d'ajustement (khi2, Anderson-Darling et autres). A partir de ces ajustements on peut calculer la durée de vie résiduelle des composants ou systèmes étudiés.

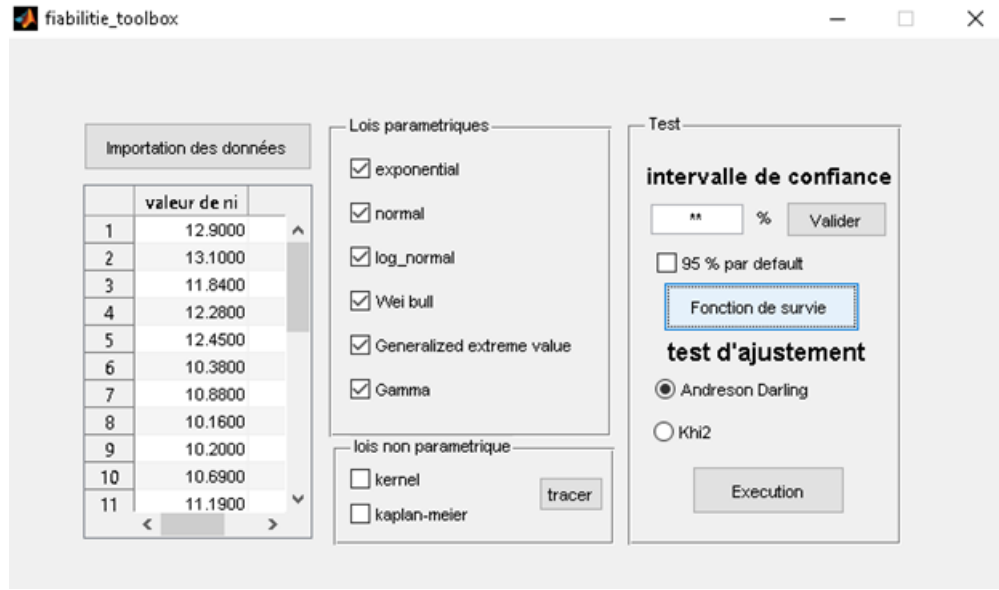


Figure III.2 L'interface graphique de la toolbox de fiabilité.

La toolbox est intégrée dans une interface graphique de Matlab (GUI) permettant de lire les données et de visualiser les courbes des différentes fonctions recherchées.

III.3.1 Menu de la toolbox

1- Zone d'importation des données

La toolbox permet d'introduire les données de durée de vie des composants ou des systèmes étudiés, et cela par différentes possibilités. La saisie peut se faire directement sur l'interface graphique dans le tableau figurant sur la GUI, ou bien importer les données à partir de fichiers. Plusieurs formats de fichiers sont pris en compte par la toolbox :

- Fichier texte '*.txt'
- Fichier de données matlab '*.mat'
- Fichier excel '*.xls' ou '*.ods'

2- Zone de sélection des Lois de probabilité

Cette zone de l'interface graphique permet de sélectionner les lois de probabilité qui vont être utilisées dans l'analyse fiabiliste. Les lois présentées sont celles qui sont le plus souvent utilisées dans le domaine de la maintenance. La toolbox permet aussi d'effectuer une analyse non paramétrique en proposant deux lois, celle du kernel et celle de Kaplan-Meier.

3- Zone test d'ajustement

Le choix de loi de probabilité, qui s'ajuste au mieux aux données de durée de vie, est effectué sur la base d'un test d'ajustement avec un niveau de confiance prédéfini à 95% ou ajuster selon le

besoin. Deux tests sont proposés dans la toolbox : le test de Khi-deux et le test de Anderson-Darling, ensuite cette zone contient l'affectation d'option « fonction de survie » qu'il approuve l'observation graphique de la fonction de survie des lois distributions avec les données empiriques dans l'intervalle de confiance précisé (95%) avant l'exécution du test.

III.3.2 Interface graphique des résultats

L'exécution du test d'ajustement permet de représenter les résultats obtenus pour les différentes lois de probabilité sélectionnées sous forme d'histogrammes **figure III.3**. Ce qui permet de trouver la loi qui s'ajuste au mieux aux données introduites de durée de vie. Néanmoins, l'utilisateur a toujours la possibilité d'imposer la loi de probabilité de son choix qui servira par la suite à l'analyse fiabiliste. La **figure III.4** montre un exemple où le résultat du test est très proche pour les lois sélectionnées.

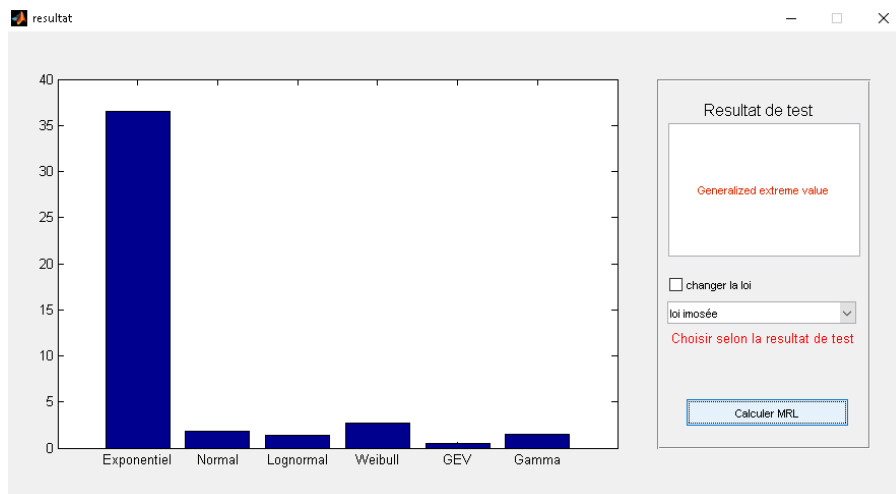


Figure III.3 Exemple de représentation des résultats du test d'ajustement.

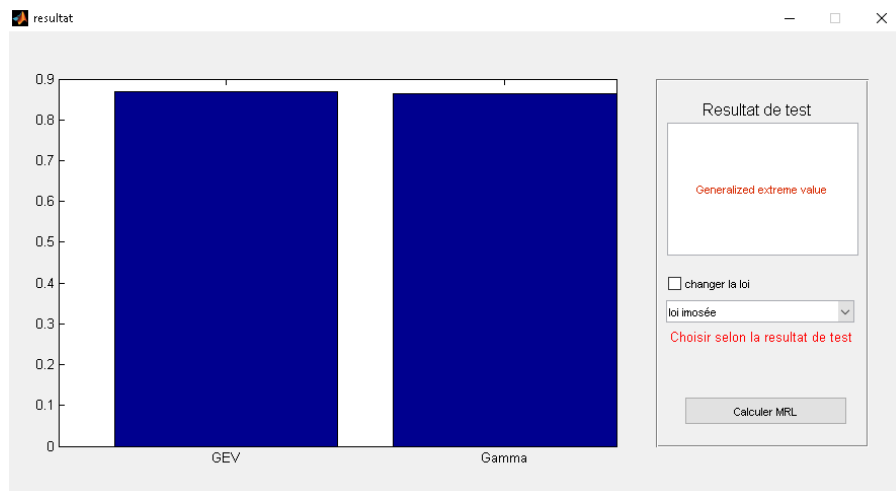


Figure.4 exemple de résultats comparables du test d'ajustement

Chapitre III: Présentation de la toolbox et application

L'interface graphique contient une zone dans laquelle les paramètres de la loi imposée par l'utilisateur peuvent être introduits **figure III.5**. D'un autre côté, les différentes fonctions de fiabilité issues de la loi de distribution sélectionnée peuvent être représentées, à savoir :

- PDF : fonction de densité de probabilité.
- CDF : fonction de densité de probabilité cumulative (fonction de répartition).
- Survie : fonction de survie (fiabilité).
- Hazard : fonction de Hazard (fonction de taux de défaillance).

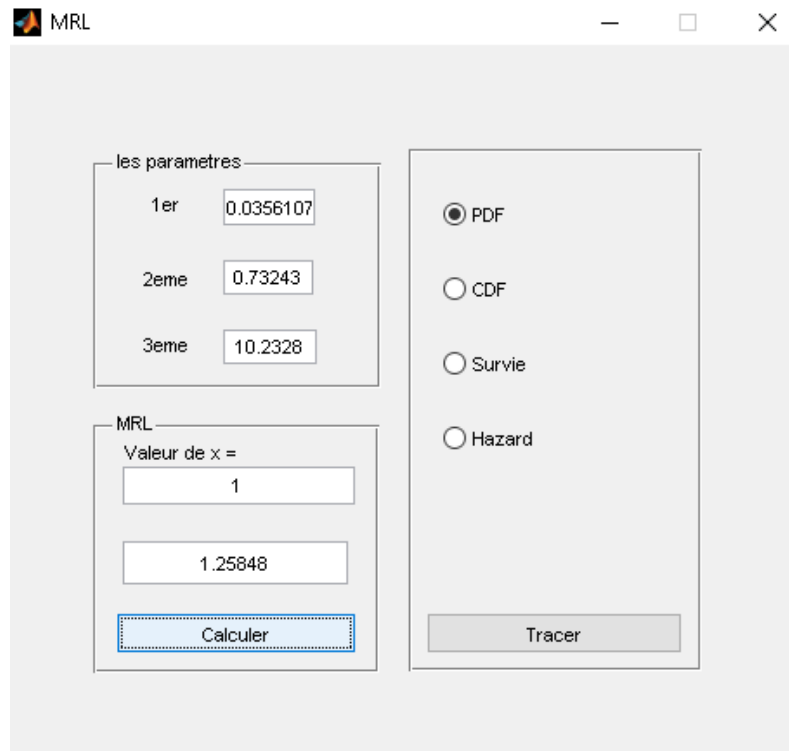


Figure III.5 Interface graphique pour le calcul de la MRL

III.4 Application de la toolbox dans le cas de l'usure des outils de coupe

Tous les outils de coupe subissent le phénomène d'usure lors de leur utilisation. La longévité d'une arête de coupe est de l'ordre d'une quinzaine de minutes pour les outils modernes. La durée de vie d'un outil de coupe est définie par la norme ISO 3685 (1993) comme étant le temps de coupe total d'un outil nécessaire pour atteindre un critère de durée de vie spécifique.

Les temps d'atteinte du seuil d'usure de plaquettes d'usinage en tournage sont présentés dans le **tableau III.1**. Ces points ont été obtenus par interpolation linéaire entre la mesure de l'usure en dépouille avant le seuil et la mesure de l'usure en dépouille après dépassement du seuil [18].

N	T (min)	N	T (min)	N	T (min)
1	12.9	11	11.19	21	10.39
2	13.1	12	10.36	22	9.77
3	11.84	13	11.08	23	11.21
4	12.28	14	10.1	24	9.87
5	12.45	15	9.05	25	10.16
6	10.38	16	10.18	26	11.22
7	10.88	17	10.11	27	10.12
8	10.16	18	10.3	28	10.17
9	10.2	19	10.48	29	9.64
10	10.69	20	9.59		

Tableau III.1 Temps de défaillance obtenus pour 29 plaquettes de coupe identiques.

III.4.1 Analyse de fiabilité

A partir des données du **tableau III.1** nous allons réaliser une analyse de fiabilité de cet outil de coupe et établir la courbe donnant l'évolution de la durée de vie résiduelle moyenne (MRL). Les données de durée de vie sont introduites dans la toolbox sous forme de fichier Excel et toutes les lois de distribution de probabilité seront sélectionnées pour cette analyse **figure III.6**. Le test d'ajustement portera sur les deux types de test proposés dans la toolbox, avec un intervalle de confiance de 95% (par défaut).

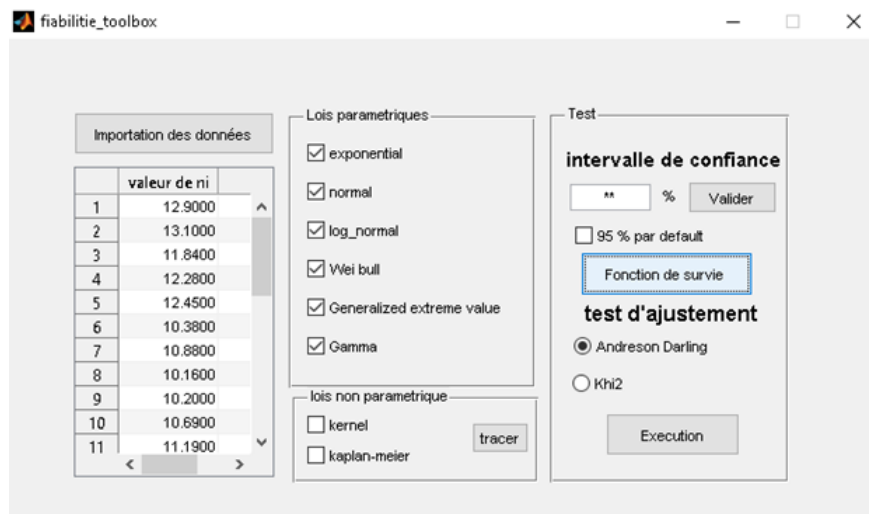


Figure III.6 Introduction des données.

La **figure III.7** représente les fonctions de survie établies selon les différentes lois de distribution sélectionnées, ainsi que la fonction de survie empirique.

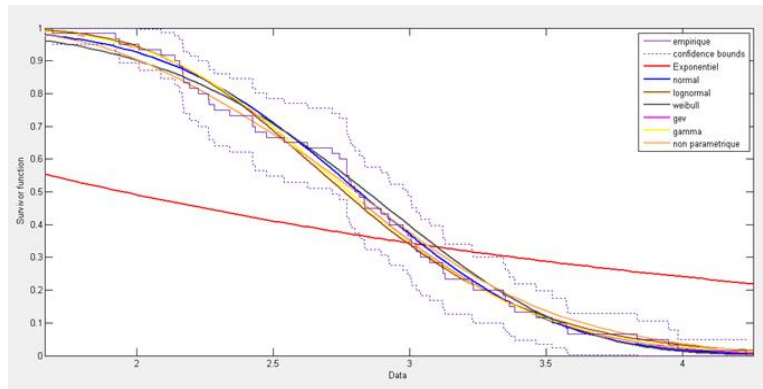


Figure III.7 Représentation des fonctions de survie

III.4.2 Analyse non paramétrique

La toolbox permet d'effectuer une analyse de fiabilité non paramétrique des durées de vie de l'outil de coupe. La **figure III.8** présente les fonctions : PDF, CDF, fonction de survie et de Hazard cumulative, obtenues par l'utilisation de l'estimateur de noyau (Kernel).

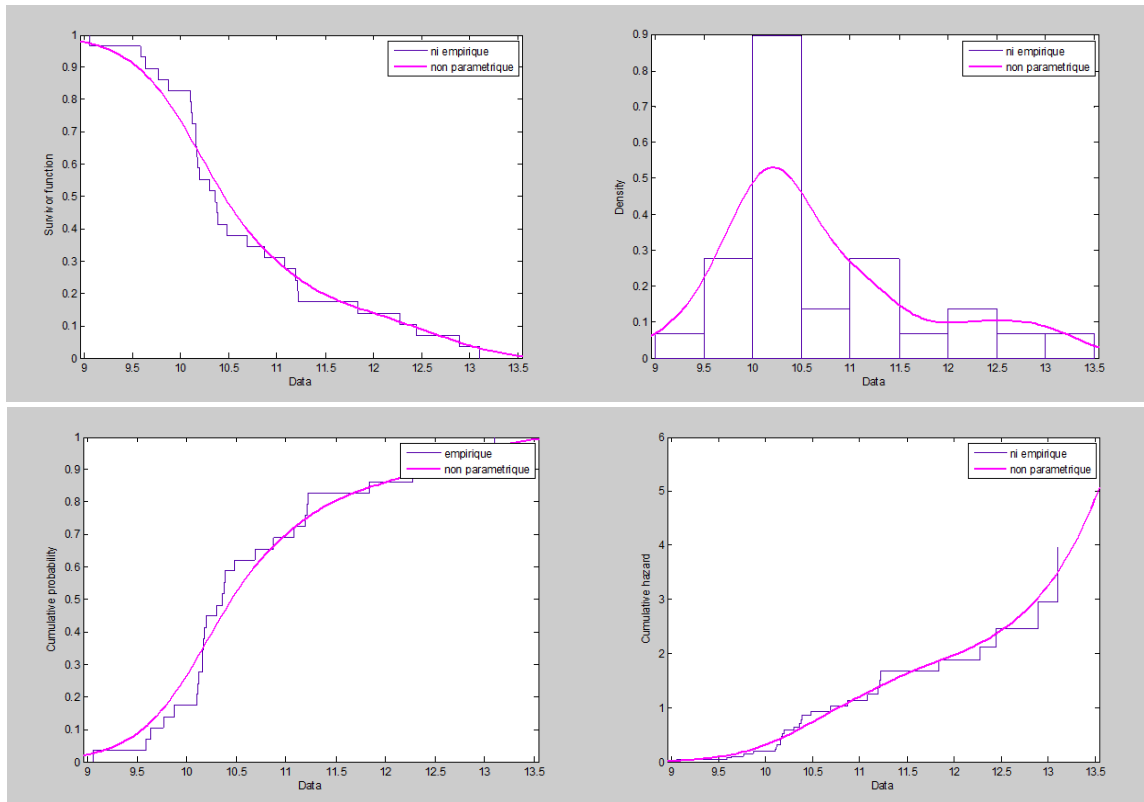


Figure III.8 Représentation des fonctions de fiabilité non paramétrique par noyau (Kernel).

III.4.3 Analyse paramétrique

La **figure III-9-a** représente les résultats du test d'Anderson-Darling obtenus pour les différentes lois de distribution utilisées, où on note que la loi de distribution de type GEV est la mieux adaptée pour décrire les données de durées de vie de l'outil de coupe. Le test de Khi-deux est également utilisé bien qu'il soit juste à titre indicatif puisque le nombre de données est inférieur à 50.

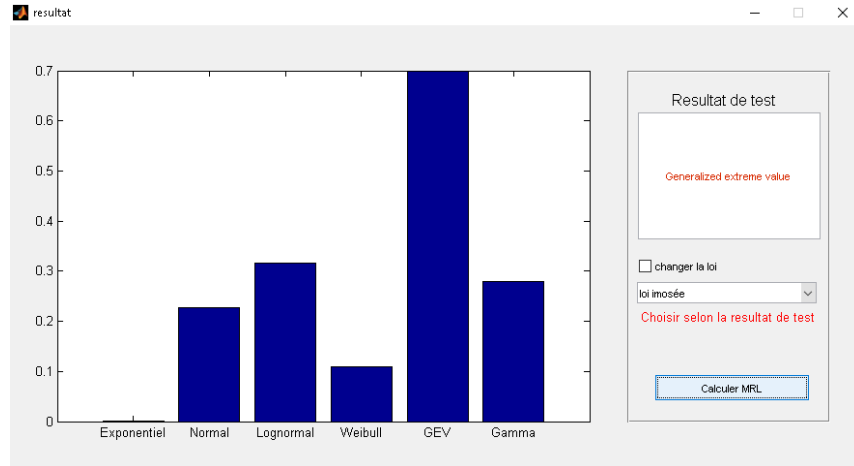


Figure III.9-a Résultats du test d'ajustement d'Anderson-Darling.

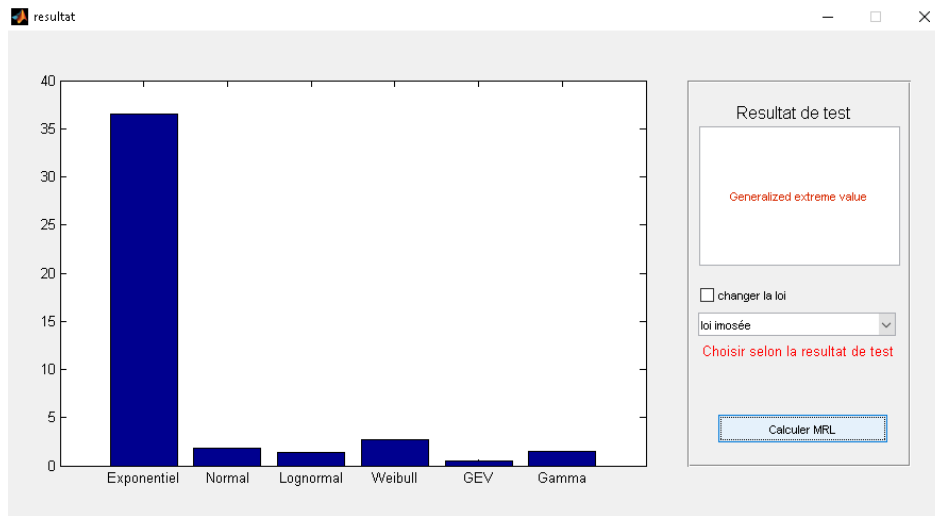


Figure III.9-b Résultats du test d'ajustement de Khi-deux.

Une fois le choix de la loi de distribution est fixé, l'interface graphique suivante, **figure III-10**, nous permet d'avoir :

- Les valeurs des paramètres de la loi GEV.

Chapitre III: Présentation de la toolbox et application

- La valeur de la durée de vie résiduelle (MRL) à partir d'un instant donné.
- Les courbes des fonctions **figure III-11** : de densité de probabilité (PDF), de survie (fiabilité), de densité de probabilité cumulée (CDF) et de Hazard cumulé.

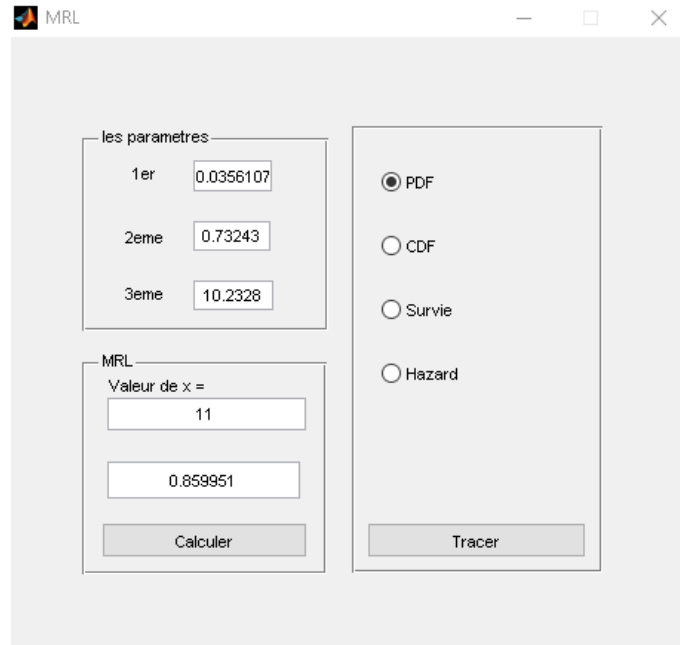


Figure III.10 Interface graphique pour la représentation des fonctions de fiabilité.

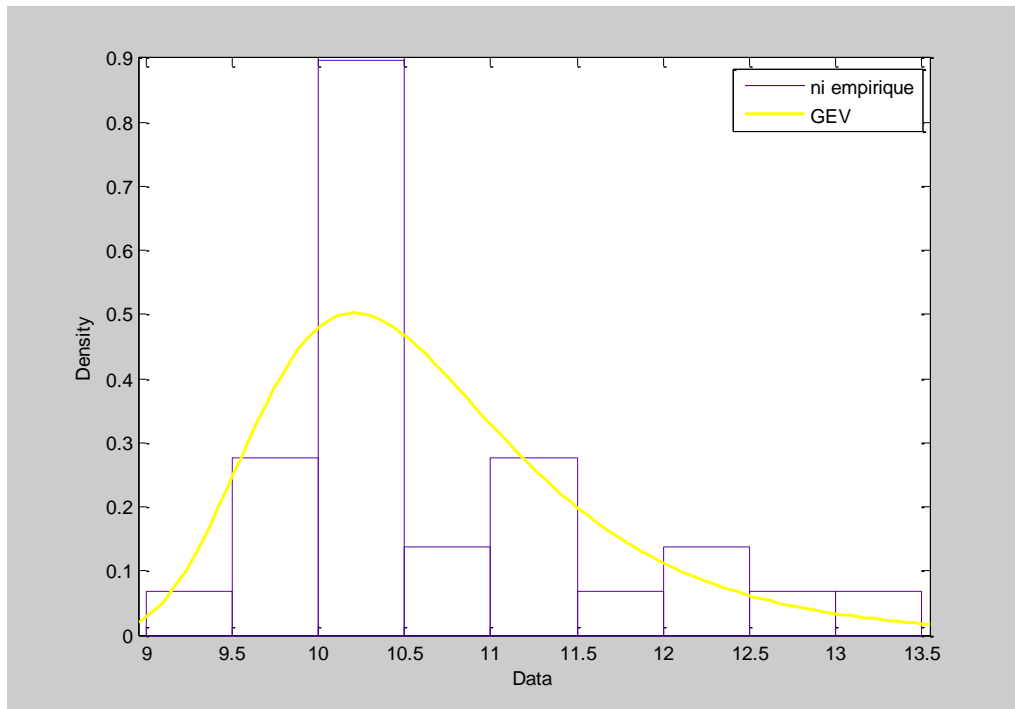


Figure III.11-a Représentation graphique de la PDF.

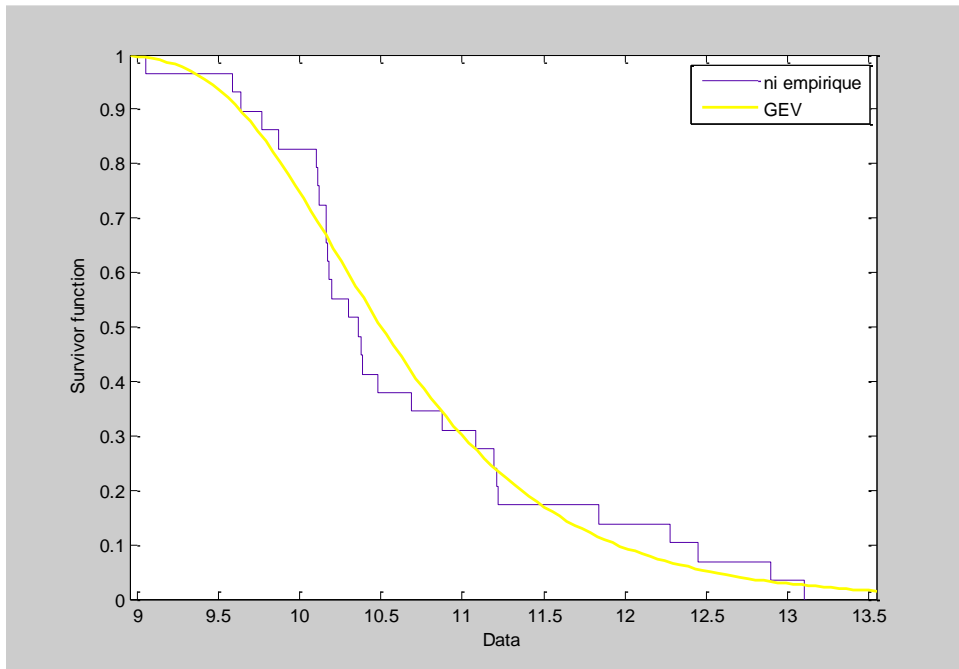


Figure III.11-b Représentation graphique de fonction de survie.

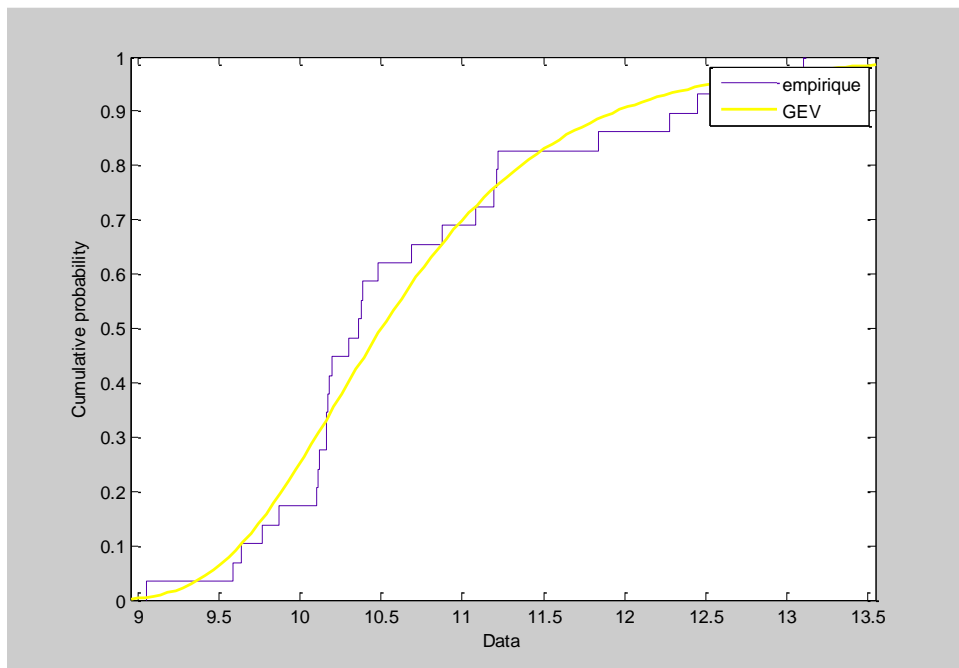


Figure III.11-c Représentation graphique de la fonction CDF.

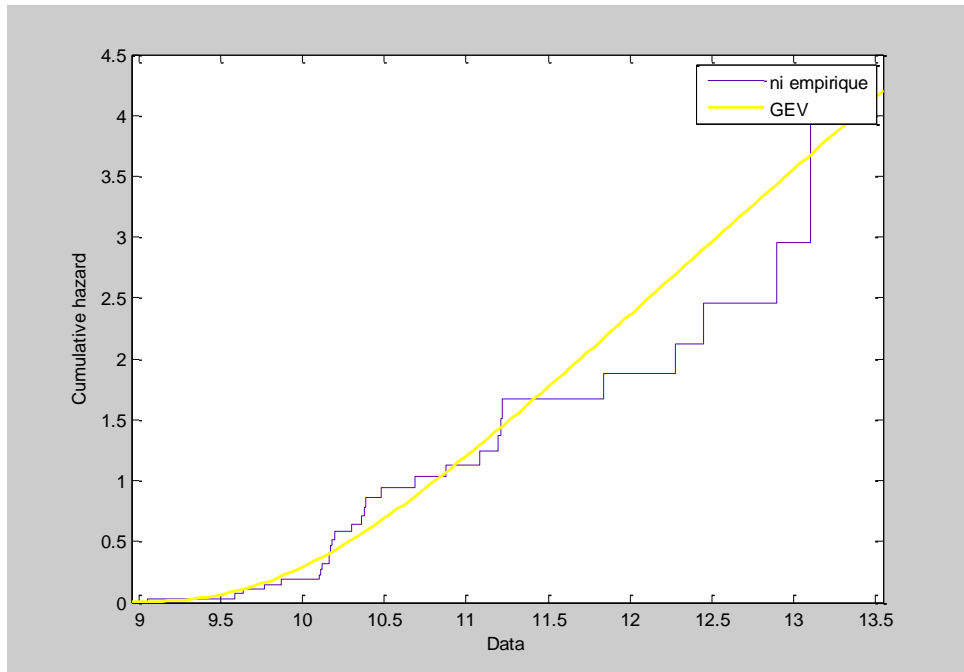


Figure III.11-d Représentation graphique de fonction de Hazard cumulative.

La durée de vie résiduelle peut être obtenue également **figure III-12** ce qui permet d'estimer pour un moment donné le temps d'usage restant avant usure de l'outil de coupe.

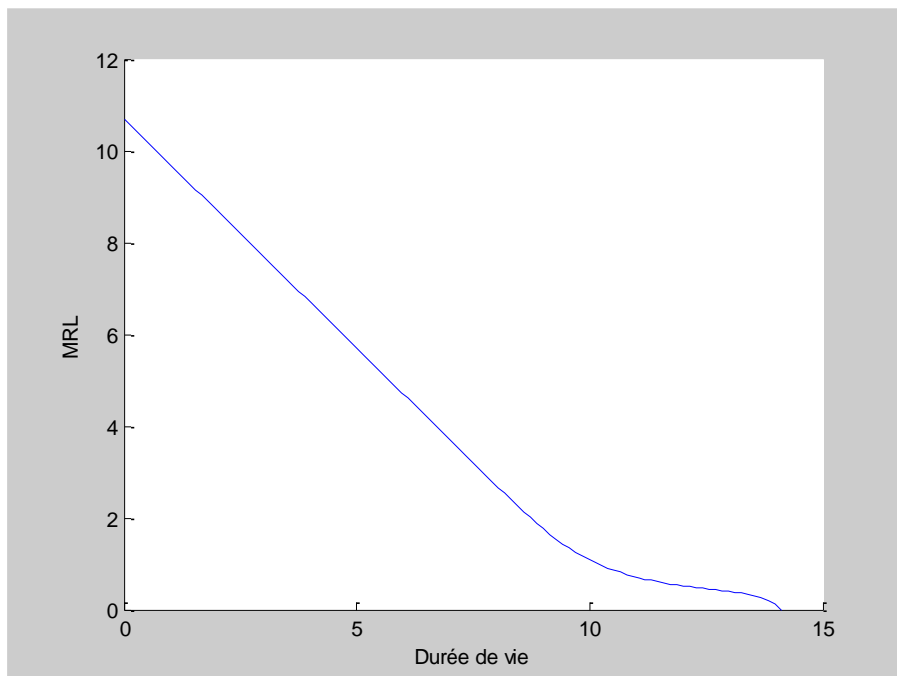


Figure III.12 Représentation de la durée de vie résiduelle.

III.5 Conclusion

Ce chapitre présente un aperçu global de la toolbox d'analyse de fiabilité. L'interface graphique développée permet une utilisation assez simple sans avoir forcément des connaissances de programmation sous Matlab. Le cas étudié de l'usure de l'outil de coupe montre les possibilités qu'offre la toolbox et son aptitude à estimer la durée de vie résiduelle de l'outil.

Conclusion générale

La fiabilité industrielle est définie comme la science des défaillances. Elle est un élément clé de l'analyse des risques et garantit la maîtrise des processus. Le développement d'outils informatiques pour assurer cette fonction est devenu plus que nécessaire. La toolbox développée dans le cadre de notre travail tente d'apporter une aide au personnel de la maintenance et offre un outil indispensable pour l'estimation de la durée résiduelle des composants et systèmes en service. La toolbox est conçue sous forme d'une interface graphique ne nécessitant aucune connaissance de langage de programmation. Il suffit juste d'introduire les données de durées de vie et la toolbox s'en charge de trouver la loi de distribution de probabilité qui s'ajuste au mieux aux données introduites. Les fonctions de survie, de défaillance, de hasard cumulé et de durée de vie résiduelle sont tracées par la suite.

L'application de la toolbox dans le cas de l'usure des plaquettes des outils de coupe montre l'intérêt du développement d'une telle toolbox. Notons que plus la base de données des durées de vie est importante, meilleur est l'estimation de la MRL. Il suffit donc d'actualiser à chaque fois le fichier de données pour que la toolbox soit plus efficace.

Références bibliographiques

- [1] Blischke, W.R. and Murthy, D.N.P. (2000). Reliability Modeling, Prediction, and Optimization. New York, NY: Wiley.
- [2] Elsayed, E. (2012). Reliability Engineering, 2e. Hoboken, NJ: Wiley.
- [3] O'Connor, P.D.T. (2012). Practical Reliability Engineering, 4e. West Sussex, UK: Wiley.
- [4] Cai, Z., Sun, S., Si, S., and Yannou, B. (2011). Identifying product failure rate based on a conditional Bayesian network classifier. *Expert Systems with Applications* 38 (5): 5036–5043.
- [5] Gupta, R.C. and Bradley, D.M. (2003). Representing the mean residual life in terms of the failure rate. *Mathematical and Computer Modelling* 37 (12/13): 1271–1280.
- [6] Claude PELLEGRIN. Fondements de la Décision de Maintenance. Réalisé en P.A.O. par STDI - Z. A. Route de Couterne - 53110 Lassay-les-Châteaux Imprimé en France. - JOUVE, 18. rue Saint-Denis. 75001 PARIS N°250116Y. - Dépôt légal Octobre 1997.
- [7] Oates, A.S. and Lin, M.H. (2009). Electromigration failure distributions of Cu/Low-k dual-damascene vias: impact of the critical current density and a new reliability extrapolation methodology. *IEEE Transactions on Device and Materials Reliability* 9 (2): 244–254.
- [8] Filippi, R.G. Wang, P.-C., Brendler, A., and Lloyd, J.R. (2010). Implications of a threshold failure time and void nucleation on electromigration of copper interconnects. *Journal of Applied Physics* 107 (103709): 1–7. doi: 10.1063/1.3357161.
- [9] Ross, S. (1998). A First Course in Probability, 5e. Upper Saddle River, NJ: Prentice Hall.
- [10] M. Idris Tioguim et M. Dimitri Delcaillau (7 juin 2018) Modélisation d'extrêmes de séries temporelles : une étude empirique. Master 1 d'actuariat – EURIA Bureau d'Etudes M1. Euro-Institut d'actuariat EURIA.
- [11] [Enligne] <https://nobelis.eu/photis/Estimat/densite.html> .
- [12] Anderson, T.W., Darling, D.A.: Asymptotic theory of certain goodness of fit criteria based on stochastic processes. *Ann. Math. Stat.* 23, 193–212 (1952).
- [13] Anderson, T.W., Darling, D.A.: A test of goodness of fit. *J. Am. Stat. Assoc.* 49, 765–769 (1954).
- [14] Renée Veysseyre Aide-mémoire Statistique Et probabilités pour l'ingénieur 2^{ème} édition. Dunod, Paris, 2001, 2006 ISBN 2 10 049994 7.
- [15] Statistics and Machine Learning Toolbox™ User's Guide.
-

[16] [Enligne] <https://fr.mathworks.com/discovery/matlab-gui.html> .

[17] [Enligne] <https://briot-jerome.developpez.com/matlab/tutoriels/introduction-programmation-interfaces-graphiques/> .

[18] Maintenance prévisionnelle d'équipements industriels basée sur la modélisation, l'estimation et la simulation de lois de dégradation, C. LETOT, thèse de doctorat 2013

