



HAL
open science

Planification des essais accélérés : optimisation, robustesse et analyse

Seyyedeh Zohreh Fatemi

► **To cite this version:**

Seyyedeh Zohreh Fatemi. Planification des essais accélérés : optimisation, robustesse et analyse. Autre. Université d'Angers, 2012. Français. NNT : . tel-01004379

HAL Id: tel-01004379

<https://theses.hal.science/tel-01004379>

Submitted on 11 Jun 2014

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Planification des essais acc l r s : optimisation, robustesse et analyse

Th se de Doctorat

Sp cialit  : Sciences de l'ing nieur

 COLE DOCTORALE STIM

Pr sent e et soutenue publiquement par

Seyyedeh Zohreh Fatemi

le 20 D cembre 2012

  l'ISTIA - Universit  d'Angers

Devant le jury ci-dessous :

| | | | |
|--------------|---|--|-----------------------------------|
| Pr sident | : | | |
| Rapporteurs | : | Patrick Lyonnet, Professeur | ENISE-Universit  Jean Monnet |
| | | L o Gerville-R ache, Ma tre de conf rences HdR | Universit  Bordeaux 2 |
| Examineurs : | | Christian B s, Professeur | Universit  Paul Sabatier-Toulouse |
| | | David Bigaud, Professeur | Universit  d'Angers |
| | | Fabrice Gu rin, Professeur | ISTIA-Universit  d'Angers |
| | | Laurent Saintis, Ma tre de conf rences | ISTIA-Universit  d'Angers |

Directeurs : Fabrice Gu rin et Laurent Saintis

Laboratoire : LABORATOIRE EN S RET  DE FONCTIONNEMENT, QUALIT  ET ORGANISATION
(UPRES EA 3858). 62, avenue Notre Dame du Lac, F-49000 ANGERS

PLANIFICATION DES ESSAIS ACCÉLÉRÉS :
OPTIMISATION, ROBUSTESSE ET ANALYSE

*Accelerated testing planning : optimization, robustness
and analysis*

Seyyede Zohreh Fatemi



Université d'Angers

Seyyedeh Zohreh FATEMI

Planification des essais accélérés : optimisation, robustesse et analyse

??+138 p.

Ce document a été préparé avec L^AT_EX₂_ε et la classe `these-IRIN` version 0.92 de l'association de jeunes chercheurs en informatique LOGIN, Université de Nantes. La classe `these-IRIN` est disponible à l'adresse :
<http://www.sciences.univ-nantes.fr/info/Login/>

Impression : ThÃse_vf.tex - 30/10/2012 - 10:31

Révision pour la classe : \$Id: these-IRIN.cls,v 1.3 2000/11/19 18:30:42 fred Exp

Résumé

La qualification d'un produit, lors des phases de développement, est une étape importante dans un projet qui vérifie que les performances et la fiabilité atteignent les objectifs. Les essais de qualification sont souvent coûteux en temps et en nombre de produits testés. Les essais accélérés consistent à soumettre des unités à des niveaux de stress plus élevés qu'en condition d'exploitation afin de réduire le temps d'apparition des défaillances. Ils permettent de construire plus rapidement la fonction de fiabilité à partir d'un modèle appropriée reliant durée de vie et stress. De plus, un plan d'essai doit être construit, précisant les paramètres du plan (niveaux de stress, allocation de l'échantillon) pour trouver le meilleur compromis entre le coût d'essai et la qualité d'estimation. L'objectif de la thèse a été de définir une méthodologie de construction de plans d'essai accéléré optimaux et robustes. Nous avons donc développé un cadre général basé sur la minimisation du coût global et une approche bayésienne. Les distributions a priori, issues de la connaissance des paramètres de fiabilité et du modèle d'accélération, sont utilisées dans l'inférence bayésienne et dans une simulation de Monte Carlo d'exploration des fiabilités possibles. Le plan optimal et robuste est obtenu à partir de méthodes d'optimisation (Surface de réponse, Algorithmes génétiques). Enfin, une méthodologie de suivi des essais est développée en observant la pertinence des résultats par rapport aux informations a priori à partir d'un facteur de similitude. Il permet de vérifier si la décision quant à la qualification peut être prise plus rapidement, ou d'optimiser le plan en cours de réalisation.

Mots-clés : Fiabilité, Essais accélérés, optimisation, estimation bayésienne, coût des essais, coût d'exploitation, robustesse

Abstract

Qualification of a product during the development phases is an important step in a project. It permits to verify that the performances and reliability achieve the objectives. The qualification tests are often costly in time and number of products tested. Accelerated tests consist of submitting units to stress levels higher than operating condition to reduce the time to failures and allow building faster the reliability function. This requires the determination or development of appropriate life-stress relationship model. In addition, a test plan should be built, accurate the design parameters (number of test levels, stress levels, sample allocation at each stress level ...) to find the best compromise between cost test and quality estimation. The objective of the thesis is to define a methodology to obtain optimal and robust accelerated test plans. So, we have developed a general framework based on the minimization of a global cost function and a Bayesian approach. The prior distributions, from the available knowledge on the parameters of reliability and acceleration model, are used in Bayesian inference and in a Monte Carlo simulation to explore possible reliabilities. The optimal and robust plan is obtained by optimization methods (Response surface and Genetic algorithms). Finally, a methodology of monitoring during test realization is developed. The relevance of the results in relation to prior information is studied using a similarity factor between observed and a priori data. It allows verifying whether the decision on the qualification can be taken more quickly in presence of more reliable product than expected or optimizing the plan under realization.

Keywords: accelerated test, reliability, optimisation, Bayesian estimation, testing cost, operation cost, robustness

Remerciements

Je souhaite remercier en premier lieu mon directeur de thèse, le professeur Fabrice Guérin, pour avoir accepté de diriger ma thèse. Je lui suis également reconnaissant pour sa disponibilité, ses qualités pédagogiques et scientifiques. J'ai beaucoup appris à ses côtés. J'en profite pour lui exprimer ici ma plus profonde gratitude.

J'adresse mes remerciements les plus chaleureux à mon autre co-encadrant Laurent Saintis, maître de conférences à Université Angers, pour tous les précieux conseils qu'il m'a donnés, pour la confiance qu'il m'a témoignée, son enthousiasme, sa disponibilité et le partage de ses nombreuses connaissances lors de nos nombreuses discussions.

Je remercie très sincèrement le professeur Abdessamad Kobi, directeur du Laboratoire en Sécurité de fonctionnement, Qualité et Organisation (LASQUO), de l'Université d'Angers de m'accepter et de me permettre d'utiliser le laboratoire.

Je voudrais remercier les rapporteurs de cette thèse professeur Patrick Lyonnet, ENISE, et M. Léo Gerville-Réache, maître de conférences HDR à université Bordeaux, pour l'intérêt qu'ils ont porté à mon travail.

Je tiens également à remercier les professeurs David Bigaud et Christian Bès pour avoir accepté d'examiner mon travail et de participer au jury.

Enfin, je souhaite remercier ma famille, Je pense tout d'abord à mes parents pour leur soutien au cours de ces années d'études et sans lesquels je n'en serai pas là aujourd'hui. Il m'est impossible de trouver des mots pour dire à quel point je suis fier d'eux, et à quel point je les aime. Je suis reconnaissant à ma mère pour ses efforts et pour m'encourager à continuer mes études. Elle est toujours dans mon cœur et dans mon esprit. Je la loue et je l'adore.

Je tiens à remercier mon fils Ali qu'il est toute ma vie, pour son excellente compréhension. Mes derniers remerciements vont à mes sœurs et mon frère pour leur soutien constant, j'associe à ces remerciements ma sœur Dr. Fatemeh Fatemi pour sa gentillesse et sa loyauté et spécialement ma sœur Dr. Sedigheh Fatemi en tant que consultant et un exceptionnel compagnon dans ma vie. Pour conclure, je souhaite bien évidemment remercier mon mari Saeid pour son soutien, sa patience et sa tolérance dans tous ces quatre ans.

Table des matières

| | |
|---|-----------|
| Table des matières | 9 |
| Introduction générale | 11 |
| Rappels sur la Fiabilité | 17 |
| 2.1 Sûreté de Fonctionnement | 17 |
| 2.1.1 Le concept FMDS | 17 |
| 2.1.2 Grandeurs associées | 19 |
| 2.2 La fiabilité | 20 |
| 2.2.1 La fiabilité tout au long du cycle de vie | 21 |
| 2.2.2 Mesures des performances | 22 |
| 2.3 Estimation de la fiabilité par une approche paramétrique | 26 |
| 2.3.1 Le modèle standard de durée de vie | 27 |
| 2.3.2 Les estimations de paramètres | 28 |
| 2.3.3 Estimation des caractéristiques de fiabilité | 32 |
| 2.4 Conclusion | 33 |
| Estimation de la fiabilité par les essais accélérés | 35 |
| 3.1 Les essais accélérés | 36 |
| 3.1.1 Introduction | 36 |
| 3.1.2 Le plan d'essais | 36 |
| 3.2 Le modèle standard de vie accélérée (SVA) | 39 |
| 3.2.1 Principe | 40 |
| 3.2.2 Lois d'accélération courantes | 42 |
| 3.2.3 La construction d'un modèle paramétrique | 43 |
| 3.3 Application numérique sur le modèle de durée de vie des roulements à billes | 46 |
| 3.4 Conclusion | 52 |
| Estimation bayésienne | 53 |
| 4.1 Introduction | 54 |
| 4.2 L'inférence Bayésienne dans le cadre de modèles SVA | 56 |
| 4.2.1 Application de l'inférence Bayésienne au modèle SVA paramétrique | 57 |
| 4.3 L'information a priori | 63 |
| 4.3.1 Sources d'information | 63 |
| 4.3.2 Le choix de la loi a priori | 65 |
| 4.3.3 Construction de la loi a priori | 69 |
| 4.4 Application numérique d'estimation bayésienne d'un modèle SVA | 71 |
| 4.4.1 A priori cohérent et peu informatif | 72 |
| 4.4.2 A priori cohérent et très informatif | 73 |
| 4.4.3 A priori très informatif et incohérent | 74 |
| 4.5 Conclusion | 75 |

| | |
|--|------------|
| Planification optimisée et robuste d'essais accélérés dans un contexte de qualification | 77 |
| 5.1 Introduction | 78 |
| 5.2 Problématique | 78 |
| 5.3 Principe de l'optimisation des plans d'essai | 79 |
| 5.3.1 Hypothèses générales | 79 |
| 5.3.2 Critère de planification et formulation du problème | 80 |
| 5.3.3 Plan d'essai pour un seul stress | 81 |
| 5.4 Analyse des méthodes existantes de construction de plans optimaux | 83 |
| 5.5 Proposition d'un plan d'essai optimal et robuste | 84 |
| 5.5.1 Définition de l'objectif de fiabilité | 86 |
| 5.5.2 Plan d'essai à optimiser | 86 |
| 5.5.3 Connaissance a priori sur le produit | 87 |
| 5.5.4 Définition de la fonction coût | 87 |
| 5.5.5 Formulation du problème d'optimisation | 88 |
| 5.6 Méthodes d'optimisation d'un plan d'essai | 88 |
| 5.6.1 Optimisation à l'aide d'une surface de réponse | 89 |
| 5.6.2 Exemple numérique pour un plan de qualification de roulements à billes | 90 |
| 5.6.3 Recherche de l'optimum à l'aide des algorithmes génétiques | 98 |
| 5.6.4 Exemple numérique pour un plan de qualification | 103 |
| 5.7 Conclusion | 105 |
| Conduite d'un plan d'essais accélérés | 107 |
| 6.1 Introduction | 108 |
| 6.2 Compatibilité de l'a priori avec les données | 108 |
| 6.2.1 Le principe de calcul d'un facteur de compatibilité normalisé | 111 |
| 6.2.2 L'utilisation du facteur de compatibilité comme facteur de pondération de l'a priori | 112 |
| 6.2.3 Application au modèle SVA | 116 |
| 6.3 Analyse finale ou périodique des résultats d'essai | 118 |
| 6.4 Analyse périodique des résultats et ajustement du plan d'essai | 118 |
| 6.5 Exemple numérique de l'analyse périodique des résultats d'essai | 119 |
| 6.5.1 Stratégie de conduite de l'essai | 120 |
| 6.5.2 Cas 1 : Acceptation du test à la première étape | 121 |
| 6.5.3 Cas 2 : Acceptation du test à la deuxième étape | 122 |
| 6.5.4 Cas 3 : Refus du test à la troisième étape | 122 |
| 6.5.5 Cas 4 : Acceptation du test à la troisième étape | 124 |
| 6.6 Conclusion | 124 |
| Conclusions et Perspectives | 127 |
| Bibliographie | 131 |

Introduction générale

Plus les caractéristiques d'un produit ou d'un système sont appréhendées tôt dans son cycle de vie, moins les risques financiers ou reliés à la sûreté des installations dus à la non réalisation des performances attendues sont élevés. Dans un contexte d'exigences de systèmes de plus en plus fiables et sûrs, et de durées de garanties croissantes, il est impératif de vérifier le plus tôt possible que les performances des systèmes sont conformes au cahier des charges.

Pour ce faire, l'idéal est de procéder à des séries d'essais sur des prototypes quand ils existent. Le problème est l'investissement en temps et en quantité de matériel important demandé car les matériels étant de plus en plus fiables, l'observation de défaillances est de moins en moins probable. Les industriels ne peuvent plus se permettre de tels coûts financiers. A l'extrême, certains systèmes se fabriquent à l'unité, ce qui rend les politiques d'essai difficiles.

Ainsi, cette problématique a été la source, pour la communauté scientifique, de nombreuses voies de recherche. Celles-ci sont basées principalement sur la modélisation stochastique des apparitions des défaillances au cours du temps et sur l'estimation statistique des paramètres des modèles à partir des résultats d'essai. Les premiers travaux ont consisté à étudier les essais réalisés en conditions nominales (on reproduit les conditions normales d'utilisation du produit), provoquant des temps d'essai importants, et avec des tailles d'échantillon élevées. Par la suite des travaux ont été menés pour réduire les tailles d'échantillon et temps d'essai :

- P. Sander, R. Badoux , H. Procaccia, C. Clarotti, A. Lannoy, J. Ringler, ... ([Sander and Badoux, 1991a], [Procaccia et al., 1992a], [Procaccia and Morilhat, 1996a], [Clarotti, 1998a], [Ringler, 1979a], [Lannoy and Procaccia, 1994a]) ont proposé d'intégrer toute la connaissance disponible sur la fiabilité des systèmes testés dans les plans d'essais, à l'aide des statistiques bayésiennes, permettant de réduire le nombre de produits à tester.
- W. Nelson, E. Elsayed, O'Connor, H. Caruso, Kececioglu, P. Hoang, V. Bagdonavicius, M. Nikulin, ... ([Nelson, 1990b], [Shyur et al., 1999b],[O'Connor, 2003a], [Caruso and Dasgupta, 1998], [kececioglu, 1944], [Hoang, 2003a], [kececioglu, 1944], [Bagdonavicius and Nikulin, 2002], ...) ont proposé d'utiliser les essais accélérés (sévérisation des conditions d'essai permettant de provoquer les défaillances plus rapidement) pour estimer la fiabilité d'un produit et les essais bayésiens
- S. Voiculescu, W. Meeker, ... ([Voiculescu and Guerin, 2005], [Zhang and Meeker, 2006]) ont étudié l'estimation bayésienne des modèles de vie accélérée utilisés pour l'analyse des résultats d'essais accélérés.

Aujourd'hui, ces techniques sont souvent utilisées par les industriels pour évaluer la fiabilité d'un équipement. Ainsi, les essais accélérés sont couramment utilisés pour la qualification afin de démontrer que l'exigence de fiabilité est tenue avec un certain niveau de confiance.

La figure 5.1 montre comment peut être définie une exigence de fiabilité.

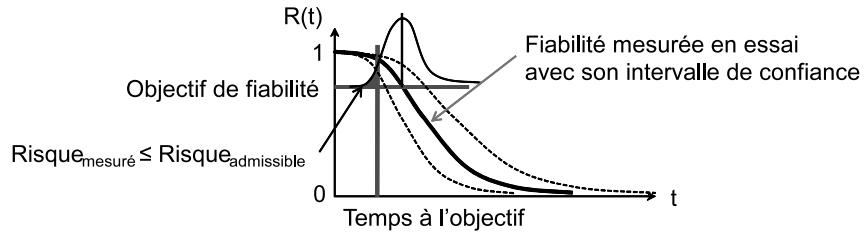


Figure 1.1 – Exigence de fiabilité

Ces exigences de qualification sont fixées dans les conditions opérationnelles d'utilisation de l'équipement caractérisées par un profil de mission plus ou moins complexe. Le profil de mission peut être constitué de deux composantes :

- Phases d'utilisation : cycle de vie du produit composé de plusieurs phases pendant lesquelles le produit rend service (cf. 1.2),

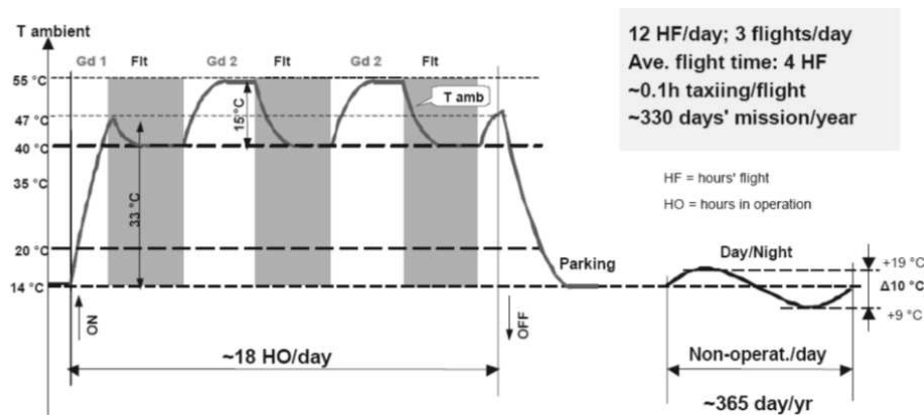


Figure 1.2 – Exemple d'un profil de mission avec plusieurs phases de vie

- Environnement de fonctionnement (climatique, mécanique,) comme l'illustre la figure 1.3

| Constraint | | Thermal and Humidity | | | Thermal cycling | | | | Mechanical | Chemical | | | Induced | |
|--------------------|-----------------------|----------------------|--------------------------|----------------------|-----------------|--------------------------|------------------------|---|-------------------------|------------------|----------------------|---------------------|-------------------------|----------------|
| Phase name | Calendar time (hours) | On/Off | Ambient temperature (°C) | Rate of humidity (%) | ΔT (°C) | Number of cycles (/year) | Cycle duration (hours) | Maximum temperature during cycling (°C) | Random vibrations (Gms) | Saline pollution | Industrial pollution | Area of application | Protection level | IT application |
| Ground-operating-1 | 797 | On | 47 | 30 | 33 | 330 | 2 | 47 | - | Low | Urban | Uninhabitable | Non-hermetically sealed | 4.8 |
| Ground-operating-2 | 1193 | On | 55 | 30 | 15 | 647 | 1.5 | 55 | - | Low | Urban | Uninhabitable | Non-hermetically sealed | 2.0 |
| Ground-taxing | 84 | On | 47 | 5 | - | - | - | - | 5 | Low | Negligible | Uninhabitable | Non-hermetically sealed | 1.2 |
| Flight-operating | 4083 | On | 40 | 5 | - | - | - | - | 0.6 | Low | Negligible | Uninhabitable | Non-hermetically sealed | 1.1 |
| Ground-dormant | 2603 | Off | 14 | 70 | 10 | 108 | 24 | 19 | - | Low | Urban | Uninhabitable | Non-hermetically sealed | 3.3 |

Figure 1.3 – Exemple d’un profil de mission avec plusieurs phases de vie et de des environnements spécifiques

Aussi, il est important de définir des essais de qualification intégrant ces profils de mission complexes. Les différents environnements (climatique, électrique, mécanique, ...) intègrent des stress (température, humidité, cyclage thermique, vibration,) pouvant solliciter des modes de défaillance indépendants (cf. table 1.1). Ainsi, il est possible de décomposer l’essai de qualification en une série d’essais indépendants avec des stress de sévèrisation spécifiques.

| Stress | Effets |
|--------------------|---|
| Température | Oxydation : dégradation des performances électriques Electromigration : dégradation des performances électriques Fluage : dégradation des performances électriques et mécaniques Interdiffusion : dégradation des performances électriques et mécaniques |
| Température cyclée | Fatigue thermique par dilatation différentielle : rupture de connexion |
| Humidité | corrosion (accélérée avec la température élevée) |
| Tension | isolation Dégradation des contacts |
| Intensité | Oxydation : dégradation des performances électriques Electromigration : dégradation des performances électriques Champs magnétiques :interférence |
| Vibration | Fatigue mécanique (fissuration, rupture,) |

Table 1.1 – Influence des principaux stress sur les mécanismes de défaillance en électronique

Ainsi, la construction du plan d’essai consiste à définir la série d’essais accélérés à réaliser en considérant que la fiabilité globale $R_{globale}$ dépend des fiabilités associées aux différents stress et phases de vie. Dans le cas où nous avons plusieurs modes de défaillances indépendants (associés des stress différents), il est usuel de procéder à des essais accélérés distincts. Lors de la qualification (fin de la phase de développement avant de lancer la production), il est très rare de disposer d’une taille d’échantillon importante et l’ingénieur est confronté à un problème de répartition des produits à tester entre les différents essais. Il s’agit ensuite pour l’ingénieur de définir les plans d’essais accélérés (combien de niveaux d’essai, allocation des produits à tester à chaque niveau, temps d’essai, ...) garantissant une bonne qualité d’estimation. Plusieurs méthodes de construction de plan optimal d’essais accélérés ont été proposées afin de maîtriser la qualité

d'estimation. La plupart de ces méthodes sont basées sur la minimisation de la variance d'estimation, afin de garantir une précision du résultat. Ces méthodes nécessitent de connaître la fiabilité potentielle de l'équipement à tester et les paramètres du modèle de vie accélérée. Cette caractérisation est souvent réalisée de manière déterministe. Pourtant, il est communément admis que les prévisions de fiabilité sont soumises à une grande incertitude. De plus, l'industriel doit avant tout maîtriser le risque pris lors de la qualification (relié à la qualité d'estimation) et les coûts associés :

- Coût d'essai augmentant avec la qualité d'estimation (diminution de la variance)
- Coût d'exploitation augmentant avec le risque pris lors de la qualification (dégradation de la qualité d'estimation) pouvant provoquer un nombre de défaillances supérieur à l'objectif.

C'est ainsi que nous proposons de construire un plan optimal et robuste permettant de prendre en compte les considérations citées précédemment :

- minimisation d'une fonction coût basée sur approche globale économique (essai et exploitation),
- prise en compte du critère de qualification intégrant l'objectif de fiabilité contractuel,
- intégration des incertitudes sur les prévisions de fiabilité et des paramètres du modèle de vie accélérée,
- estimation bayésienne des paramètres de modèle.

Finalement, on obtient un plan qui permet de garantir la prise de décision quant à la qualification. Toutefois, ce plan permet de couvrir un domaine assez vaste de la fiabilité de l'équipement et des paramètres du modèle de vie accélérée. Aussi, lors de la réalisation de l'essai, nous proposons d'évaluer le critère de qualification afin de savoir s'il est possible de prendre une décision plus rapidement et d'agir directement sur les paramètres de plan. Pour ce faire, nous étudierons la cohérence des résultats d'essai avec les connaissances a priori qui ont permis de construire le plan en utilisant le facteur de pondération développé par S. Voiculescu ([Voiculescu and Guérin, 2009]) pour les modèles de vie accélérée.

Ainsi, le document est organisé de la manière suivante :

- Le premier chapitre rappelle brièvement les définitions de quelques indicateurs de fiabilité et techniques d'estimation statistique de paramètres de modèles qui seront utilisés dans la suite du document ([Villemeur, 1988a], [Cocozza-Thivent, 1997], [Pages and Gondran, 1980], [Afnor, 1988]).
- Le deuxième présente les essais accélérés en insistant sur le modèle Standard de Vie Accélérée (SVA) paramétrique ([Nelson, 1990b], [Crowe and Feinberg, 2001a], [ASTE, 1993], [O'Connor, 2003a], [Ligeron and et M. Neff, 1984]).
- Le troisième introduit l'inférence bayésienne et son application aux modèles SVA paramétriques. ([Crowe and Feinberg, 2001a], [McLean, 2000], [kececioglu and Sun, 1999])
- Le quatrième présente la problématique de la construction de plans optimaux. Les limites des méthodes actuelles seront précisées afin de dégager les objectifs pour développer une nouvelle approche. Ainsi, nous présenterons une nouvelle approche permettant d'obtenir un plan optimal et robuste en précisant :
 - la fonction Coût basée sur une approche économique intégrant la métrique de fiabilité contractuelle,
 - la procédure de construction de plan utilisant la simulation de Monte Carlo et l'inférence bayésienne,
 - les méthodes d'optimisation permettant de rechercher le plan optimal
- Le cinquième chapitre présente différentes approche pour l'analyse des résultats d'essai au

cours de la réalisation du plan optimal et robuste. Ces approches consistent à évaluer le critère de qualification pour vérifier si la décision peut être prise plus rapidement car la fiabilité réelle le permet.

Enfin le document se termine par des conclusions et des perspectives.

Rappels sur la Fiabilité

Ce chapitre a pour objectif d'introduire brièvement quelques notions de fiabilité et les techniques d'estimation statistique de paramètres de modèle.

2.1 Sûreté de Fonctionnement

La **Sûreté de Fonctionnement** (*SdF*) est l'aptitude d'une entité à satisfaire une ou plusieurs fonctions requises dans des conditions données, la propriété qui permet aux utilisateurs du système de placer une confiance justifiée dans le service qu'il leur délivre, la capacité de maintenir le niveau de performance dans le temps.

La sûreté de fonctionnement désigne à la fois un ensemble de moyens et un ensemble de résultats produits par ces moyens [Mortureux, 2001]:

- les méthodes et outils pour caractériser et maîtriser les effets des aléas, des pannes, des erreurs [Villemeur, 1988b], [Pagès and Gondran, 1980], [O'Connor, 2002];
- les caractéristiques des systèmes pour exprimer la conformité de leurs comportements et leurs actions à long terme [Pagès and Gondran, 1980], [Procaccia and Morilhat, 1996b], [Pham, 2003].

Au sens large, la SdF est considérée comme la science des défaillances et des pannes. Différents auteurs définissent la SdF comme :

- la science des défaillances [Villemeur, 1988b];
- la confiance justifiée dans le service qu'il délivre [Laprie et al., 1995];
- le maintien de la qualité dans le temps [Mortureux, 2001];
- l'ensemble des propriétés qui décrivent la disponibilité et les facteurs qui conditionnent la fiabilité, la maintenabilité et la logistique de maintenance [ISO 9000].

Plus personne aujourd'hui n'accepterait d'utiliser un produit, un système dont tous les aspects liés à son usage ne soient pas optimisés sous ses quatre aspects de la Sûreté de Fonctionnement [Fournier, 1993]:

1. La **F**iabilité
2. La **D**isponibilité
3. La **M**aintenabilité
4. La **S**écurité

Dans les paragraphes suivants, nous allons définir ces quatre attributs FMDS de la SdF et les métriques associées.

2.1.1 Le concept FMDS

Le concept **FMDS** est apparu aux États-Unis dans les années 1960 dans l'industrie du spatial et de l'armement.

C'est un ensemble d'outils et de méthodes qui vont permettre sur toutes les phases de la vie d'un produit de s'assurer qu'il va réaliser la (les) mission(s) pour laquelle (lesquelles) il a

été conçu dans des conditions de Fiabilité, de Maintenabilité, de Disponibilité et de Sécurité définies. Ainsi, la SdF doit être prise en compte tout au long du cycle de développement du produit. Les quatre attributs de la FMDS sont définis de la manière suivante:

La fiabilité

La fiabilité (reliability en anglais) est l'aptitude d'un dispositif à accomplir une fonction requise dans des conditions données pour une période de temps donnée [Chapouille, 1972]. Elle se caractérise par la probabilité $R(t)$ que l'entité E accomplisse ces fonctions, dans les conditions données pendant l'intervalle de temps $[0, t]$, sachant que l'entité n'est pas en panne à l'instant 0.

$$R(t) = P[E \text{ non défaillante sur } [0, t]]$$

La maintenabilité

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

$$M(t) = P[E \text{ est réparée sur } [0, t]]$$

La disponibilité

La disponibilité (availability en anglais) est l'aptitude d'une entité à être en état d'accomplir les fonctions requises dans les conditions données et à un instant donné. Elle se caractérise par la probabilité $A(t)$ que l'entité E soit en état, à l'instant t , d'accomplir les fonctions requises dans des conditions données :

$$A(t) = P[E \text{ non défaillante à l'instant } t]$$

La sécurité

La sécurité (safety en anglais) est l'aptitude d'un produit à respecter, pendant toutes les phases de sa vie, un niveau acceptable de risques d'accidents susceptibles d'occasionner une agression du personnel, une dégradation majeure du produit ou de son environnement.

$$S(t) = P[E \text{ évite les événements critiques ou catastrophiques sur } [0, t]]$$

Nous admettrons par la suite que le temps est la variable principale dont dépend la fiabilité. Pour certains appareils, il peut être plus approprié de considérer une autre variable : le nombre de cycles d'ouverture-fermeture pour un relais, le nombre de tours pour un moteur, le nombre de kilomètres pour une voiture, voire une composition de celles-ci, etc.

La fiabilité est une composante essentielle de la sûreté de fonctionnement. Elle participe à la disponibilité d'un équipement. Afin d'envisager une étude de sûreté de fonctionnement exhaustive, il sera nécessaire de réaliser des études complémentaires dans les domaines de la maintenabilité, de la sécurité et des calculs probabilistes de la disponibilité.

2.1.2 Grandeurs associées

Des métriques, basées sur l'estimation de moyennes de temps, peuvent être utilisées pour caractériser la sûreté de fonctionnement d'un système [Villemeur, 1988b] :

- **MTTF** : La durée moyenne de fonctionnement d'une entité avant la première défaillance (anglais Mean Time To Failure)

$$MTTF = \frac{\text{temps cumulés de fonctionnement jusqu'à la 1}^{\text{ère}} \text{ défaillance}}{\text{nombre de 1}^{\text{ères}} \text{ défaillances}} \quad (2.1)$$

- **MUT** : Durée moyenne de fonctionnement après réparation (anglais Mean Up Time)

$$MUT = \frac{\text{temps cumulés de fonctionnement après la 1}^{\text{ère}} \text{ défaillance}}{\text{nombre d'intervalles entre 2 défaillances successives}} \quad (2.2)$$

- **MDT** : Durée moyenne d'indisponibilité après défaillance (anglais Mean Down Time)

$$MDT = \frac{\text{temps cumulés d'arrêts}}{\text{nombre total de défaillances après la 1}^{\text{ère}}} \quad (2.3)$$

- **MTBF** : Durée moyenne entre deux défaillances (anglais Mean Time Between Failure)

$$MTBF = \frac{\text{temps cumulés d'exploitation après la 1}^{\text{ère}} \text{ défaillance}}{\text{nombre total de défaillances après la 1}^{\text{ère}}} \quad (2.4)$$

- **MTTR** : Durée moyenne de réparation (anglais Mean Time To Repair)

Ces durées sont représentées dans la figure (2.1.2).

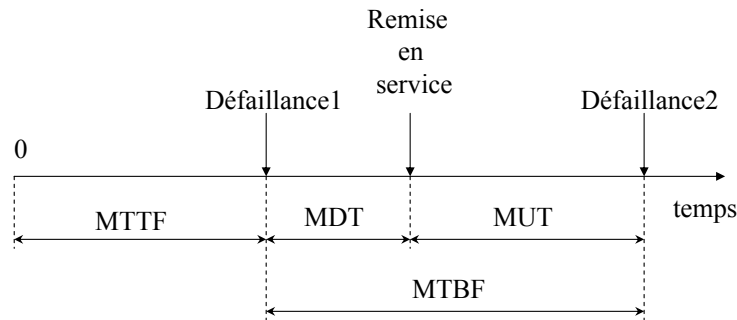


Figure 2.1 – Grandeurs associées à la sûreté de fonctionnement

La disponibilité asymptotique est donnée par :

$$\text{Disponibilité} = \frac{MUT}{MTBF} \quad (2.5)$$

Dans la suite de cette thèse, nous allons nous focaliser tout particulièrement sur la fiabilité qui est un élément essentiel de la sûreté de fonctionnement.

2.2 La fiabilité

L'analyse de la fiabilité constitue une phase indispensable dans toute l'étude de la sûreté de fonctionnement. A l'origine, la fiabilité concernait les systèmes à haute technologie et à fort potentiel de risque (nucléaire, aéronautique, spatial, ...). Aujourd'hui, la fiabilité est devenue un paramètre clef de la qualité et de l'aide à la décision dans l'étude de la plupart des composants, produits et processus "grand public": transport, énergie, bâtiments

De nombreux industriels travaillent à l'évaluation et l'amélioration de la fiabilité de leurs produits au cours de leur cycle de développement, de la conception à la mise en service (conception, fabrication et exploitation) afin de développer leurs connaissances sur le rapport Coût/Fiabilité et maîtriser les sources de défaillance.

La dégradation de la fiabilité d'un produit est directement liée aux mécanismes de dégradation de produit et/ou de ses composants. Parmi les mécanismes de défaillance nous pouvons citer :

- la fatigue,
- le fluage,
- la fissuration,
- l'usure,
- la corrosion, l'oxydation,
- l'électromigration, etc.

Les étapes nécessaires à la réalisation d'une étude complète de fiabilité pour obtenir soit des mesures de performance soit des grandeurs associées au produit sont :

1. Les objectifs de l'étude;
2. L'analyse fonctionnelle;
3. La conception (bureau d'études);
4. L'identification des risques potentiels;
5. La modélisation du système;
6. Les essais de validation sur prototypes, présérie;
7. La validation par l'analyse de retour d'expérience (REX).

On observe que jusqu'à l'étape 5 incluse, on peut parler d'une approche prévisionnelle, avant la réalisation du produit final ou de prototypes. Cette approche très utile est apte à fournir des informations pertinentes sur les caractéristiques de la fiabilité du produit en cours du développement. Elle permet rapidement et avec de ressources limitées de décider quelles sont les actions à mener pour améliorer l'équipement. Néanmoins, il est rarement possible d'intégrer les interactions entre les composants et les différents modes de défaillance qui peuvent pourtant dégrader significativement la fiabilité.

Aussi, lors de l'étape suivante, les essais de validation sur des prototypes ou des produits de présérie permettent d'étudier le comportement et les performances en simulant les conditions d'utilisation (stress, sollicitations, environnement, ...) : on parle alors d'une approche expérimentale.

La dernière étape (étape 7) permet de mesurer la fiabilité opérationnelle du produit en traitant les données issues de l'exploitation. Cette étape permet de confirmer les différentes hypothèses prises (mécanismes de défaillance influents, lois de fiabilité, ...) et de capitaliser cette

information à travers des recueils de données ou pour servir comme information a priori dans le cadre de démarches Bayésiennes pour de futures versions du produit.

Dans le cadre de la thèse, nous nous sommes intéressés à la fiabilité expérimentale. Dans les paragraphes suivants, nous présentons les différentes techniques d'essais de fiabilité et les statistiques qui peuvent être estimées à partir des résultats d'essais.

2.2.1 La fiabilité tout au long du cycle de vie

La fiabilisation d'un produit au cours du cycle de vie de développement fait appel à de nombreuses phases dans lesquelles on utilise divers méthodes et outils (voir Figure 2.2).

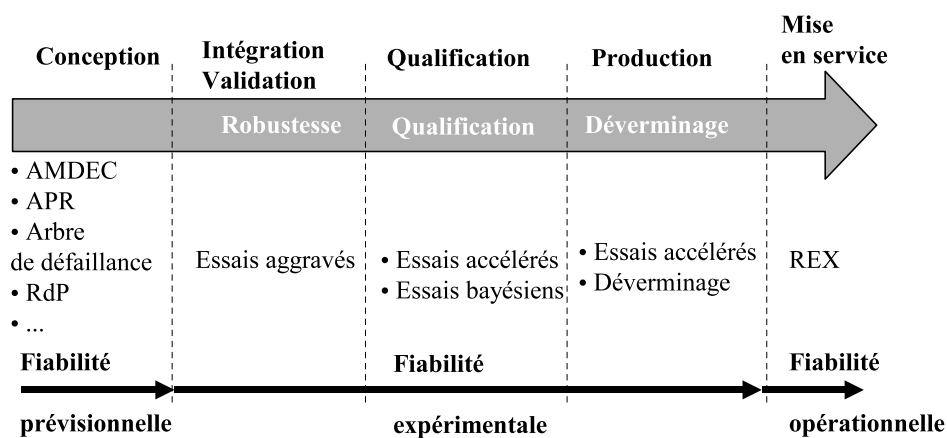


Figure 2.2 – Position des essais dans le cycle de développement d'un produit

L'ensemble des essais de fiabilité (robustesse, estimation et déverminage) contribue largement à la croissance de la fiabilité du produit au cours de son développement et de sa production (voir Figure 2.3).

Les dernières techniques d'essais développées consistent à accélérer cette croissance de fiabilité par l'utilisation d'**essais sévères** [Crowe and Feinberg, 2001b]. Ainsi, on peut citer :

- essais de robustesse : les essais HALT (Highly **Accelerated** Life Test)
- essais d'estimation : les essais accélérés ALT (**Accelerated** Life Test)
- essais de déverminage : les essais HASS (Highly **Accelerated** Stress Screen)

La cohérence entre les différentes études de fiabilité doit être gardée d'une manière constante pendant les cycles de vie d'un produit comme une partie intégrante de l'ensemble d'outils SdF.

Les estimations expérimentale et opérationnelle de la fiabilité sont basées principalement sur l'analyse et l'interprétation des instants de défaillance. La grande différence entre les deux est les moyens utilisés pour obtenir les temps de défaillance. Par contre, les méthodes et les outils d'analyse sont communs. Lorsque la fiabilité expérimentale utilise des instants de défaillance obtenus à partir d'essai sur des prototypes, la fiabilité opérationnelle utilise les instants de défaillance obtenus à partir de l'exploitation du produit.

Nous allons nous intéresser aux essais d'évaluation qui vont permettre d'évaluer des mesures de performance. Il s'agit principalement de mesures statistiques que nous allons présenter dans le paragraphe suivant.

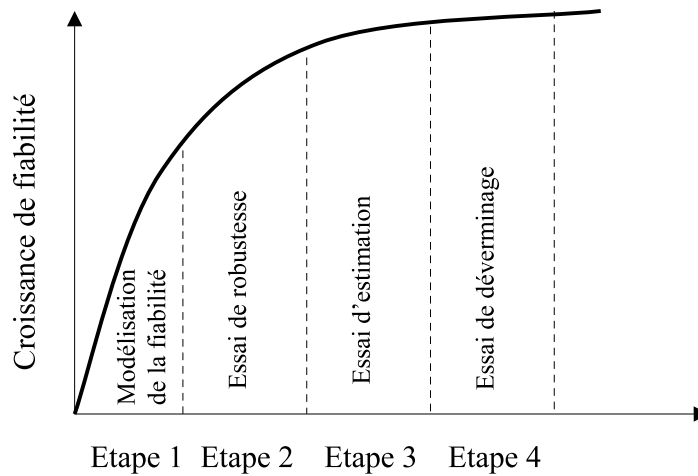


Figure 2.3 – Croissance de la fiabilité au cours du développement d'un produit

2.2.2 Mesures des performances

On définit la **défaillance** comme la cessation de l'aptitude d'une entité à accomplir une fonction requise [Villemeur, 1988b]. On va considérer T la variable aléatoire qui représente le temps écoulé entre la mise en marche d'un matériel et la première défaillance observée.

Une **variable aléatoire** est une fonction définie sur l'ensemble des résultats possibles d'une expérience aléatoire, telle qu'il soit possible de déterminer la probabilité pour qu'elle prenne une valeur donnée ou qu'elle prenne une valeur dans un intervalle donné [Tiplica, 2007].

La **théorie des probabilités** est l'étude mathématique des phénomènes caractérisés par le hasard et l'incertitude. Plus précisément, les probabilités sont utilisées pour modéliser des situations dans lesquelles une même expérience réalisée à des différents moments mais dans les mêmes conditions, peut produire des résultats différents.

Toutes les mesures des performances suivantes sont déclinées de la théorie des probabilités pour une variable aléatoire X , adaptée pour le cas concret $X = T$, et interprétée en association avec leurs sens physiques.

Fonction de fiabilité

La fonction de fiabilité ou fonction de survie $R(t)$, est définie comme l'aptitude d'une entité à accomplir une fonction requise, dans des conditions données, pendant un intervalle de temps donné.

La fiabilité à l'instant t est la probabilité qu'une entité E soit non défaillante sur la durée $[0; t]$, en supposant qu'elle n'est pas défaillante à l'instant $t = 0$.

Par extension, on appelle également fiabilité, la probabilité associée $R(t)$ définie par :

$$R(t) = P(T > t) \quad (2.6)$$

Les aspects théoriques de la fiabilité consisteront donc en une application particulière de la théorie des probabilités aux problèmes de durée de fonctionnement sans incidents qui va utiliser toutes les informations autour du sujet.

La figure 2.4 ci-dessous présente une allure de la fonction $R(t)$ en fonction du temps.

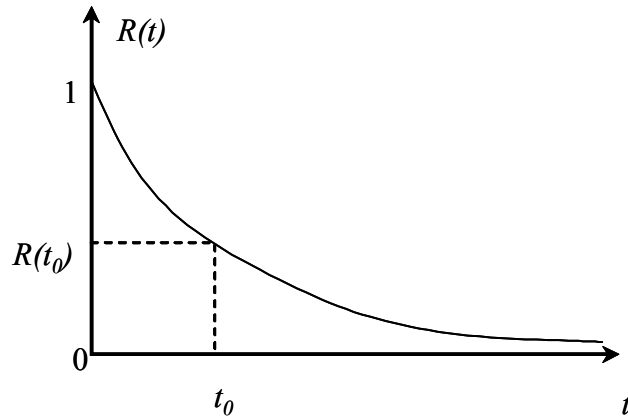


Figure 2.4 – Courbe de fiabilité

Pour compléter l'approche théorique de la notion de fiabilité, il est nécessaire de définir aussi les notions suivantes, qui sont dérivées de la théorie des probabilités.

La fonction de répartition

La fonction $F(t)$ de l'équation 2.7 représente la fonction de répartition de la variable aléatoire T . C'est la défiabilité $\overline{R}(t)$ (la probabilité de défaillance du système) ou le complément à 1 de la probabilité de fiabilité $R(t)$ définie par:

$$\begin{aligned} F(t) &= P(T \leq t) \\ F(t) &= \overline{R}(t) = 1 - R(t) \end{aligned} \quad (2.7)$$

La fonction de répartition est la primitive de la densité de probabilité $f(t)$ (voir le paragraphe suivant). La fonction de répartition a les propriétés suivantes :

- $F(t)$ est croissante;
- $F(t)$ est partout continue à droite;
- $\lim_{t \rightarrow 0} F(t) = 0$;
- $\lim_{t \rightarrow \infty} F(t) = 1$.

Densité de probabilité

Si dt est un nombre réel positif infiniment petit (pour nous un intervalle de temps), alors la probabilité d'obtenir T inclus dans l'intervalle $[t; t + dt]$ est égale à $f(t)dt$, soit:

$$P(t < T < t + dt) = F(t + dt) - F(t) \quad (2.8)$$

avec :

$$\frac{F(t + dt) - F(t)}{dt} = f(t) \quad (2.9)$$

Finalement, on obtient la relation générale de la fonction de densité :

$$f(t) = \frac{dF(t)}{dt} = -\frac{dR(t)}{dt} \quad (2.10)$$

sachant que $f(t)$ est une fonction localement intégrable pour la mesure de Lebesgue .

De même, nous pouvons ajouter les relations suivantes :

$$\begin{aligned} F(t) &= \int_0^t f(u)du \\ R(t) = 1 - F(t) &= 1 - \int_0^t f(u)du = \int_t^\infty f(u)du \end{aligned} \quad (2.11)$$

Le taux de défaillance instantané

Il est aussi possible de définir la notion de taux instantané de défaillance au temps t , notée $\lambda(t)$. La valeur $\lambda(t)dt$ représente la probabilité d'avoir une défaillance dans l'intervalle de temps $[t; t + dt]$, sachant qu'il n'y a pas eu de défaillance dans l'intervalle de temps $[0; t]$.

Ainsi, en appliquant le théorème des probabilités conditionnelles, puis le théorème des probabilités totales, $\lambda(t)$ s'écrit :

$$\begin{aligned} \lambda(t)dt &= \frac{\text{Prob(défaillant sur } [t; t + dt] \text{ sans défaillance sur } [0; t])}{\text{Prob(non défaillant sur } [0; t])} \\ \lambda(t)dt &= \frac{\text{Prob(défaillant sur } [0; t + dt]) - \text{Prob(défaillant sur } [0; t])}{\text{Prob(non défaillant sur } [0; t])} \end{aligned}$$

soit finalement

$$\lambda(t) = \frac{f(t)}{R(t)} = -\frac{1}{R(t)} \cdot \frac{dR(t)}{dt} \quad (2.12)$$

On peut aussi remarquer que :

$$\begin{aligned} \lambda(t) &= -\frac{d \ln [R(t)]}{dt} \\ \text{d'où : } \ln [R(t)] &= -\int_0^t \lambda(u)du \\ \text{et : } R(t) &= \exp \left[-\int_0^t \lambda(u)du \right] \end{aligned} \quad (2.13)$$

Au cours de la vie d'un matériel, de manière empirique, nous pouvons distinguer trois phases sur l'évolution du taux de défaillance $\lambda(t)$ au cours du temps t , à travers une courbe caractéristique dite en "baignoire" (figure 2.5).

Les trois phases sur l'évolution du taux de défaillance sont:

- la première concerne les défaillances précoces dues à des problèmes de conception (mauvais dimensionnement d'un composant, ...) ou de production (dérive d'un process de fabrication,...) ;

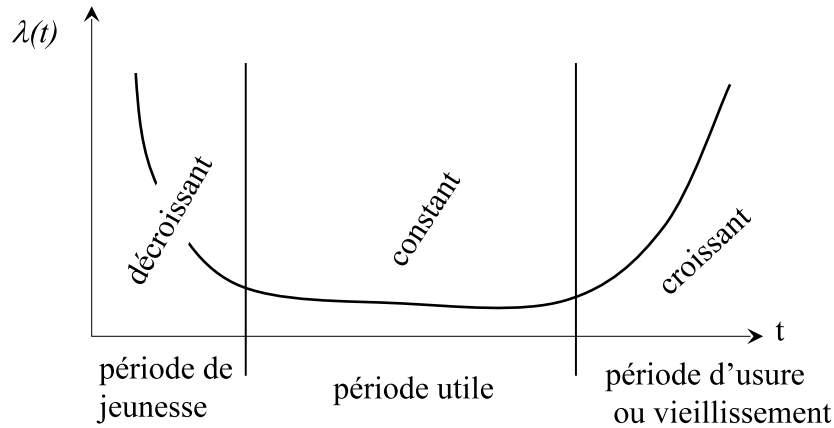


Figure 2.5 – Courbe en baignoire

- la deuxième partie, pouvant être plus ou moins importante selon le type de matériel (plus pour l'électronique et moins pour la mécanique), est caractéristique des défaillances aléatoires ;
- la troisième partie correspond aux défaillances dues à des phénomènes d'usure, de vieillissement, etc.

Le percentile

Le $P^{\text{ème}}$ percentile d'une distribution $F(t)$ est l'instant τ_P par lequel une proportion P de la population est défaillante. C'est la solution de

$$\tau_P = F^{-1}(P) \quad (2.14)$$

Les moments

On appelle moment d'ordre k de la variable T , lorsqu'il existe, le réel défini par :

$$E[T^k] = \int_{-\infty}^{+\infty} t^k f(t) dt \quad (2.15)$$

$E[X]$ est l'espérance mathématique ou moyenne. La moyenne peut être estimée directement à partir d'un échantillon de taille n par :

$$E[T] = m = \frac{\sum t_i}{n} \quad (2.16)$$

On appelle variance le réel $V[T]$, lorsqu'il existe, défini par :

$$V[T] = \int_{-\infty}^{+\infty} (t - E[T])^2 f(t) dt \quad (2.17)$$

La variance peut être également estimée par :

$$V[T] = \frac{\sum (t_i - m)^2}{n - 1} \quad (2.18)$$

$\sqrt{V[T]}$ est appelé écart type (noté $\sigma(T)$); d'où $V[T] = \sigma(T)^2$

La moyenne et l'écart type permettent de caractériser, respectivement, la tendance et la dispersion.

Mode

Le mode permet également de caractériser la tendance. C'est la valeur du caractère statistique qui apparaît le plus fréquemment. Dans le cas d'une variable continue, on peut entendre parler de classe modale qui serait la classe de plus grand effectif.

2.3 Estimation de la fiabilité par une approche paramétrique

La plupart des cas d'analyse de fiabilité utilisent comme données les instants de défaillance résultant d'essais ou de l'exploitation.

La figure 2.6 présente, des instants de défaillance t_i ainsi que des suspensions τ_i pour un échantillon de taille n . La suspension matérialise le fait que l'on ne connaît pas exactement le temps de défaillance d'un équipement. Plusieurs situations peuvent produire des suspensions :

- l'essai est arrêté avant l'apparition de toutes les défaillances (censure à droite),
- la défaillance est constatée après son apparition lors d'une inspection par exemple (censure à gauche),
- la défaillance est apparue entre deux instants d'inspection (temps de censure par intervalle)

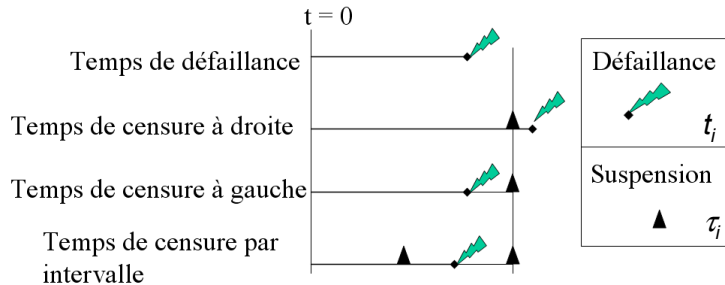


Figure 2.6 – Les types de données

A partir de ces données, des conclusions sur la loi de fiabilité peuvent être tirées. En fonction des informations dont on dispose déjà sur le produit testé, on peut faire des hypothèses plus ou moins restrictives afin d'améliorer les estimations, comme :

- l'existence ou non d'une loi paramétrique de la fonction de fiabilité
- un mode de défaillance unique,
- plusieurs modes de défaillance, etc.

Dans le cadre de la thèse, nous considérons le cas d'un seul mode de défaillance et de données censurées à droite correspondant à la majorité des cas observés lors des essais.

A partir de ces hypothèses, différentes approches sont possibles; une classification rapide en deux catégories peut être opérée selon la connaissance ou non d'une loi paramétrique pour décrire la courbe de fiabilité :

- L'approche non paramétrique : elle est basée sur aucune hypothèse sur l'existence d'une loi pour approximer la courbe de fiabilité.

- L’approche paramétrique : elle fait l’hypothèse d’une loi continue, dépendante de certains paramètres et qui approxime la courbe réelle de fiabilité.

La phase de qualification d’un produit intervient en aval des études de fiabilité. Nous nous plaçons ici dans le cas où la connaissance du produit est suffisamment forte, notamment par la connaissance d’une loi caractérisant la durée de vie. Nous présentons uniquement par la suite l’approche correspondante à cette hypothèse.

L’approche paramétrique, d’après son nom, utilise des lois paramétriques pour approximer la fonction de fiabilité. Pour ce faire, il est nécessaire de choisir la loi adaptée au mécanisme de défaillances ou au produit. Pendant des années, de nombreux chercheurs ont développé différentes lois statistiques et/ou ont associé les lois mathématiques avec le processus physique de défaillance. Dans l’analyse statistique des données, un modèle paramétrique est supposé pour la distribution inconnue ; mais le modèle est employé juste comme une approximation de la distribution inconnue. Il est alors permis de considérer que le degré de confiance d’une analyse paramétrique s’appuie sur la façon dans laquelle le modèle choisi approche la distribution inconnue.

2.3.1 Le modèle standard de durée de vie

Une forme générale de fonctions de fiabilité est décrite par le modèle standard de durée de vie. On dit que la variable aléatoire T , temps jusqu’à la première défaillance, suit une loi de fiabilité $R(t)$ si elle représente la probabilité que T soit supérieur à t (équation 2.6). La probabilité que la durée de vie soit supérieure à t , $R(t)$ est continue à droite en tout point.

On suppose que la fonction de fiabilité R appartient à une classe de fonctions qui ne dépend que des paramètres d’échelle, de forme et de décalage [Bagdonavicius et al., 2000]:

$$R(t) = R\left(\left(\frac{t - \omega}{\eta}\right)^v\right), (\eta, v > 0) \quad (2.19)$$

avec $\theta = \{\eta, v, \omega\}$ vecteur des paramètres inconnus (η , v et ω respectivement les paramètres d’échelle, de forme et de position (ou de décalage)).

Si $\omega \neq 0$, alors ce terme matérialisera la durée de vie consommée lors d’un déverminage trop agressif ($\omega > 0$) ou d’un délai de mise en service ($\omega < 0$). Dans le cas des essais, on ne retrouve pas cette situation car les équipements à tester sont produits pour l’essai, le déverminage est contrôlé et le vieillissement est considéré négligeable. Aussi, nous n’allons traiter que les cas où $\omega = 0$.

Les loi classiquement utilisées en fiabilité (Weibull, loglogistique et lognormale) appartiennent à cette famille de modèles. Dans les paragraphes suivants, nous détaillons uniquement la loi de Weibull utilisée dans les applications numériques ultérieures, ainsi qu’un de ses cas particuliers, la loi exponentielle.

La lois Weibull En considérant $R(t) = e^{-t}$ dans l’équation 2.19 on obtient la classe de fonction de fiabilité de Weibull avec :

$$R(t) = e^{-\left(\frac{t}{\eta}\right)^v} \quad (2.20)$$

et la fonction de densité de probabilité :

$$f(t) = \frac{v}{\eta} \left(\frac{t}{\eta}\right)^{v-1} e^{\left(-\frac{t}{\eta}\right)^v} \quad (2.21)$$

C'est un modèle pouvant caractériser l'ensemble de la courbe en baignoire tout au long de la vie d'un produit. Ce type de loi est plutôt utilisé pour les composants mécaniques.

Le paramètre de forme v donne des indications sur le mode des défaillances et sur l'évolution du taux de défaillances dans le temps [Lannoy, 1996] :

- $v < 1 \Rightarrow \lambda(t)$ décroît \Rightarrow période de jeunesse
- $v = 1 \Rightarrow \lambda(t)$ constant \Rightarrow indépendance du temps, loi exponentielle (paragraphe 2.3.1)
- $v > 1 \Rightarrow \lambda(t)$ croît \Rightarrow période de vieillissement ou d'usure :
 - $1.5 < v < 2.5 \rightarrow$ phénomène de fatigue
 - $3 < v < 4 \rightarrow$ phénomène d'usure (remarque : $v = 3.5 \Rightarrow$ distribution normale)

La loi exponentielle C'est une loi Weibull de paramètre $v = 1$:

$$R(t) = e^{-\lambda t}, \quad \text{avec } \lambda = \frac{1}{\eta} \quad (2.22)$$

La loi exponentielle est la plus populaire car elle est la plus simple, mais elle suppose un taux de défaillance constant. Elle ne dépend que d'un seul paramètre : le taux de défaillance λ . Cette loi est associée au processus de Poisson qui génère des évènements (défaillances) dont les temps d'occurrence t sont indépendants et distribués identiquement selon une loi exponentielle. La loi exponentielle est souvent utilisée en fiabilité pour mesurer la durée de vie des composants et des circuits électroniques. L'une des spécificités de cette loi est donnée par le fait qu'il n'y a pas d'effet mémoire. L'équation 2.19 permet aussi, pour $R(t) = \frac{1}{1+t}$ d'avoir la classe loglogistique et si $R(t) = 1 - \Phi(\ln t)$, avec Φ la fonction de répartition de la loi normale standard, on a la classe lognormale. Les méthodes d'estimation des paramètres à partir d'un nombre fini d'observations sont basées principalement sur la modélisation stochastique des apparitions des défaillances au cours du temps et sur l'estimation statistique des courbes de fiabilité à partir des résultats d'essai.

2.3.2 Les estimations de paramètres

Une fois la classe de loi définie, la démarche statistique permet d'avoir des estimations des paramètres inconnus en considérant les instants de défaillance.

L'estimation d'un paramètre, considéré comme une variable aléatoire, comprend l'estimateur ponctuel et son intervalle de confiance. Plusieurs méthodes possibles sont à la disposition de l'analyste (méthode des moments, maximum de vraisemblance, bayésienne non informative, etc.). La suite de la thèse va utiliser les méthodes du maximum de vraisemblance et bayésienne.

Pour chaque méthode, des techniques spécifiques existent pour déterminer les intervalles de confiance des estimateurs. Pour la méthode de maximum de vraisemblance, nous exploiterons l'information de Fisher en prenant une hypothèse de normalité sur les estimateurs. Pour la méthode bayésienne non informative, nous allons utiliser la construction des fonctions marginales pour caractériser complètement la variable aléatoire.

Les deux méthodes sont basées sur l'utilisation de la fonction de vraisemblance qu'il convient de préciser ici.

Les durées successives de fonctionnement de n produits identiques sont observées jusqu'à l'instant de censure t_c . Au cours de l'observation, on enregistre les instants de défaillance ($t = (t_1, \dots, t_k)$ avec k le nombre de défaillance observées).

On considère la fonction de fiabilité R de paramètres $\theta = (\theta_1, \dots, \theta_r) \in \Omega$. Alors, en utilisant $f(t, \theta)$, on définit la fonction de vraisemblance par :

$$L(t, \theta) = \left(\prod_{j=1}^k f(t_j, \theta) \right) R(t_c, \theta)^{(n-k)} \quad (2.23)$$

Les estimations par la méthode du maximum de vraisemblance (MV) : La méthode du maximum de vraisemblance consiste dans un premier temps à évaluer la vraisemblance $L(t, \theta)$ des observations, c'est-à-dire la probabilité d'avoir obtenu un certain nombre de défaillances et de censures à des instants ou des intervalles de temps donnés pour un échantillon, x .

Si $L(t, \theta)$ est dérivable et si le maximum de vraisemblance $\hat{\theta} = (\hat{\theta}_1, \dots, \hat{\theta}_r)$ existe, alors il satisfait l'équation suivante :

$$\hat{\theta} = \text{Argmax}(L(t, \theta)) \quad (2.24)$$

qui implique :

$$\left. \frac{\partial L(t_1 \dots t_n, \theta_1 \dots \theta_r)}{\partial \theta_j} \right|_{\theta_j = \hat{\theta}_j} = 0 \quad (2.25)$$

L'information de Fisher est définie comme la matrice des dérivées d'ordre deux de la log vraisemblance calculée pour $\theta = \hat{\theta}$ qui satisfait 2.25 :

$$I(\theta) = E \left[\left(\frac{\partial \log L(T, \theta)}{\partial \theta} \right)^2 \Big|_{\hat{\theta}} \right] \quad (2.26)$$

Dans l'hypothèse où la distribution de probabilité dépend de plusieurs paramètres, θ n'est plus un scalaire mais un vecteur $\theta = (\theta_1, \theta_2, \dots)$, la recherche du maximum de vraisemblance ne se résume donc pas à une seule équation mais à un système :

$$E \left[\frac{\partial}{\partial \theta_i} \log f(T; \theta) \right] = 0, \quad \forall i \quad (2.27)$$

Enfin, l'information de Fisher n'est plus définie comme une variance scalaire mais comme une matrice de covariance :

$$I(\theta_i, \theta_j) = E \left[\left(\frac{\partial}{\partial \theta_i} \log f(T; \theta) \right) \left(\frac{\partial}{\partial \theta_j} \log f(T; \theta) \right) \right] \quad (2.28)$$

En effet, le passage de l'espace des observations à l'espace des paramètres est un changement de système de coordonnées. L'inverse de cette matrice permet quant à elle de déterminer les covariances relatives aux estimations conjointes des différents paramètres à partir des observations. Le fait que tous les paramètres soient à estimer simultanément rend l'estimation plus difficile. Ce phénomène est une manifestation de ce qui est parfois appelé le *fléau de la dimension* [Bellman, 1961].

Soit $\psi(\theta)$ le vecteur d'estimateur $E[\mathbf{T}]$ avec $\mathbf{T} = (T_1, \dots, T_n)^T$.

La borne Cramér-Rao impose que la matrice de covariance pour \mathbf{T} respecte :

$$\text{cov}_\theta(\mathbf{T}) \geq \frac{\partial \psi(\theta)}{\partial \theta} [I(\theta)]^{-1} \left(\frac{\partial \psi(\theta)}{\partial \theta} \right)^T \quad (2.29)$$

avec :

- l'inégalité des matrices $A \geq B$ signifie que la matrice $A - B$ est une matrice semi définie
- $\partial \psi(\theta)/\partial \theta$ est une matrice qui a le ij élément donné par $\partial \psi_i(\theta)/\partial \theta_j$.

Si \mathbf{T} est un estimateur non biaisé de θ (i.e., $\psi(\theta) = \theta$), alors l'inégalité de Cramér-Rao se simplifie [Fisher, 1922] :

$$\text{cov}_\theta(\mathbf{T}) \geq I(\theta)^{-1} \quad (2.30)$$

Alors, pour calculer la variance de $\hat{\theta}_i$, $\hat{V}ar(\hat{\theta}_i)$, la log vraisemblance $L_i(t_i, \theta_i)$ doit être dérivable d'ordre deux. Ensuite, nous construirons la matrice symétrique de Fisher ainsi :

$$\hat{I} = \begin{bmatrix} -\frac{\partial^2 L_i}{\partial \theta_{i1}^2} & -\frac{\partial^2 L_i}{\partial \theta_{i1} \partial \theta_{i2}} & \cdots & -\frac{\partial^2 L_i}{\partial \theta_{i1} \partial \theta_{ir'}} \\ -\frac{\partial^2 L_i}{\partial \theta_{i2} \partial \theta_{i1}} & -\frac{\partial^2 L_i}{\partial \theta_{i2}^2} & \cdots & -\frac{\partial^2 L_i}{\partial \theta_{i2} \partial \theta_{ir}} \\ \cdots & \cdots & \cdots & \cdots \\ -\frac{\partial^2 L_i}{\partial \theta_{ir} \partial \theta_{i1}} & -\frac{\partial^2 L_i}{\partial \theta_{ir} \partial \theta_{i2}} & \cdots & -\frac{\partial^2 L_i}{\partial \theta_{ir}^2} \end{bmatrix} \quad (2.31)$$

La matrice inverse \hat{I}^{-1} de Fisher représente la matrice estimée de variance – covariance dénommée $\hat{\Sigma}$:

$$\hat{\Sigma} = \begin{bmatrix} Var(\hat{\theta}_{i1}) & Cov(\hat{\theta}_{i1}, \hat{\theta}_{i2}) & \cdots & Cov(\hat{\theta}_{i1}, \hat{\theta}_{ir}) \\ Cov(\hat{\theta}_{i2}, \hat{\theta}_{i1}) & Var(\hat{\theta}_{i2}) & \cdots & Cov(\hat{\theta}_{i2}, \hat{\theta}_{ir}) \\ \cdots & \cdots & \cdots & \cdots \\ Cov(\hat{\theta}_{ir}, \hat{\theta}_{i1}) & Cov(\hat{\theta}_{ir}, \hat{\theta}_{i2}) & \cdots & Var(\hat{\theta}_{ir}) \end{bmatrix} = \hat{I}^{-1} \quad (2.32)$$

Le théorème de la limite centrale est un ensemble de résultats sur la convergence faible d'une suite de variables aléatoires en probabilité.

Généralement, les évaluations de MV des paramètres sont asymptotiquement normales, signifiant que, pour des grandes tailles d'échantillon, une distribution de paramètre de la même population serait très près de la distribution normale. Ainsi, si $\hat{\theta}$ est l'estimateur de MV pour θ , dans le cas d'une distribution d'un seul paramètre, estimée à partir d'une grande taille d'échantillon, on a :

$$z = \frac{\hat{\theta} - \theta}{\sqrt{Var(\hat{\theta})}} \quad (2.33)$$

avec :

$$P(x \leq z) \rightarrow \Phi(z) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^z e^{-\frac{t^2}{2}} dt \quad (2.34)$$

Pour C_1 et C_2 respectant :

$$P(C_1 < \theta < C_2) = \alpha \tag{2.35}$$

avec α un niveau de confiance donné, l'équation 2.34 devient :

$$P\left(-w_{1-\frac{\alpha}{2}} < \frac{\hat{\theta} - \theta}{\sqrt{Var(\hat{\theta})}} < w_{1-\frac{\alpha}{2}}\right) \cong \alpha \tag{2.36}$$

avec w_α définit par $\alpha = \frac{1}{\sqrt{2\pi}} \int_{w_\alpha}^{\infty} e^{-\frac{t^2}{2}} dt$

Alors, en simplifiant l'équation 2.36, pour un certain niveau de confiance α , on a les limites de confiance :

$$\hat{\theta} - w_{1-\frac{\alpha}{2}} \sqrt{Var(\hat{\theta})} < \theta < \hat{\theta} + w_{1-\frac{\alpha}{2}} \sqrt{Var(\hat{\theta})} \tag{2.37}$$

Les limites de confiance pour le cas où θ doit être positif sont obtenues en traitant le $\log(\theta)$ comme une variable normal distribuée ; l'équation 2.37 devient :

$$\hat{\theta} e^{-\frac{w_{1-\frac{\alpha}{2}} \sqrt{Var(\hat{\theta})}}{\hat{\theta}}} < \theta < \hat{\theta} e^{\frac{w_{1-\frac{\alpha}{2}} \sqrt{Var(\hat{\theta})}}{\hat{\theta}}} \tag{2.38}$$

La méthode bayésienne non informative : Cette méthode consiste à appliquer la démarche bayésienne (qui sera très largement présentée dans le Chapitre 4) en considérant un a priori non informatif. Le théorème de Bayes s'écrit dans le cas général :

$$g(\theta|t) = \frac{L(t, \theta) \times g(\theta)}{\int_{\Omega} L(t, \theta) \times g(\theta) d\theta} \tag{2.39}$$

Les fonctions de densité $g(\theta)$ et $g(\theta|t)$ représentent respectivement l'a priori et l'a posteriori. En faisant l'hypothèse que chaque variable θ_i est indépendante, alors la distribution a priori de θ peut s'écrire comme le produit de fonctions de densité associées à chaque paramètre θ_i :

$$g(\theta) = \prod_{i=1}^r g_i(\theta_i) \tag{2.40}$$

En prenant des lois uniformes non informatives (comme définie dans la figure 2.7) les fonctions de densité deviennent indépendantes des paramètres θ_i .

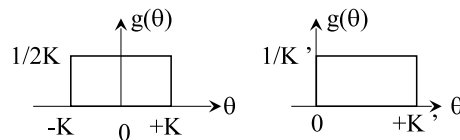


Figure 2.7 – Lois uniformes non informatives

Alors, la fonction de densité a posteriori dépend uniquement de la vraisemblance :

$$g(\theta|t) = \frac{L(t, \theta)}{\int_{\Omega} L(t, \theta) d\theta} \quad (2.41)$$

La fonction de densité $g_i(\theta_i)$ du paramètre θ_i s'écrit :

$$g_i(\theta_i) = \int_{j=1..n, j \neq i} \dots \int g(\theta|t) d\theta_j \quad (2.42)$$

Dans ce contexte, l'estimateur ponctuel est le moment initial d'ordre 1 :

$$\hat{\theta}_i = \int_{\Omega_{\theta_i}} \theta_i g_{\theta_i}(\theta_i) d\theta_i \quad (2.43)$$

La variance est définie par le moment centré d'ordre 2 :

$$V(\hat{\theta}_i) = \int_{\Omega_{\theta_i}} g_{\theta_i}(\theta_i) (\theta_i - \hat{\theta}_i)^2 d\theta_i \quad (2.44)$$

La méthode présentée nécessite de recourir à des méthodes numériques lorsque les expressions analytiques n'existent pas. Les variances sont utilisées pour faire les estimations des intervalles de confiance selon les relations 2.37 et 2.38. Il est possible également de définir l'intervalle de confiance bilatérale symétrique au niveau de confiance $1 - \alpha$ tel que :

$$prob(\theta_i^{min} \leq \theta_i \leq \theta_i^{max}) = 1 - \alpha \quad (2.45)$$

avec θ_i^{min} et θ_i^{max} solutions de :

$$\int_{D(\theta_i)}^{\theta_i^{min}} g_i(\theta_i) d\theta_i = \alpha/2 \quad (2.46)$$

$$\int_{D(\theta_i)}^{\theta_i^{max}} g_i(\theta_i) d\theta_i = 1 - \alpha/2 \quad (2.47)$$

2.3.3 Estimation des caractéristiques de fiabilité

A partir des estimations des paramètres θ de la loi de fiabilité, il est possible d'évaluer des caractéristiques particulières de fiabilité comme par exemple : une probabilité de défaillance pour un temps donné, le MTTF, le L10, ... Ces évaluations doivent s'accompagner d'une estimation d'un intervalle de confiance intégrant la variabilité sur chaque paramètre θ_i . Pour ce faire, nous utiliserons l'algèbre des variables aléatoires.

Soit la caractéristique y de fiabilité à estimer reliée à la fonction de fiabilité, de paramètre θ , par la relation $f(\theta)$ (par exemple, la durée L10 est définie par : $L10 = f(\theta) = R^{-1}(0, 1, \theta)$)

L'estimateur ponctuel de la grandeur y est donné par :

$$\hat{y} \approx f(\hat{\theta}) \quad (2.48)$$

L'intervalle de confiance sur y peut être estimé à partir des relations 2.37 et 2.38 en utilisant la variance $V(\hat{y})$ définie par :

$$V(\hat{y}) \approx \sum_{i=1}^r \sum_{j=1}^r \left(\frac{\partial f(\theta)}{\partial \theta_i} \right)_{\hat{\theta}} \left(\frac{\partial f(\theta)}{\partial \theta_j} \right)_{\hat{\theta}} COV(\hat{\theta}_i, \hat{\theta}_j) \quad (2.49)$$

Pour des cas complexes, il est possible de faire appel à des méthodes numériques ou de recourir à la technique du Bootstrap.

2.4 Conclusion

Dans les cas d'essais classiques, l'estimation de la fiabilité nécessite de tester de nombreux individus pendant longtemps afin de précipiter suffisamment de défaillances. Ces conditions sont nécessaires pour obtenir une bonne qualité d'estimation. Toutefois, ces conditions sont devenues incompatibles avec les contraintes industrielles. Aussi, une alternative est apparue avec les essais accélérés qui permettent de palier à la contrainte de temps. Dans le chapitre suivant, nous présentons les essais accélérés.

Estimation de la fiabilité par les essais accélérés

3.1 Les essais accélérés

3.1.1 Introduction

Les essais accélérés se composent d'une variété de techniques d'essai pour diminuer la durée de vie des produits ou accélérer les processus de dégradation. Ils sont utilisés pour obtenir plus vite des informations concernant la vie du produit ou concernant le rapport performance/dégradation dans le temps. Les unités testées sont employées plus fréquemment que d'habitude ou sont soumises à des niveaux de sollicitation plus élevés (par exemple l'augmentation de la température, de la tension électrique, de l'humidité, etc.), pour obtenir la loi de fiabilité ou d'autres caractéristiques de fiabilité (taux de défaillance, temps de défaillance, etc.) des produits (systèmes ou composants) dans des délais plus courts. Pour cela, les niveaux de stress subis par le produit sont augmentés afin d'obtenir des données de vie plus rapidement, qui seront utilisées pour estimer la fiabilité dans les conditions normales de fonctionnement. Les résultats sont employés, par un modèle statistique approprié basé sur la physique de défaillance des composants, pour faire des prévisions de durées de vie du produit soumis aux conditions normales d'utilisation.

Les essais accélérés peuvent s'appliquer à toutes les catégories de matériels en employant différents types de contraintes (mécaniques, électriques, climatiques ...) les plus adaptées vis-à-vis des modes de défaillance attendus sur ces matériels /

- **Contraintes mécaniques** : torsion, flexion, flexion rotative, chocs mécaniques, vibrations, vibrations acoustiques, traction/compression, etc. La fatigue est le mode de dégradation le plus communément utilisé pour les composants mécaniques qui sont testés sur des pots vibrants,
- **Contraintes électriques** : cela inclue la tension, l'intensité du courant, la fréquence, choc électrique ... La tension est le stress électrique le plus communément utilisé puisqu'il génère rapidement des défaillances,
- **Contraintes climatiques (ou environnementales)** : la température et les cycles thermiques sont les stress les plus couramment utilisés. Il est nécessaire d'appliquer des niveaux appropriés pour conserver les modes de défaillance d'origine. D'autres stress peuvent être appliqués comme les ultraviolets, le brouillard salin, la poussière, l'humidité, etc.
- **Contraintes spécifiques au matériel.**

Ces différentes contraintes peuvent être appliquées combinées ou non aux produits. Dans le cadre général, les stress X peuvent varier au cours du temps et être multidimensionnels [Nelson, 1990a] :

$$X = X(\tau), \tau \geq 0 \quad (3.1)$$

3.1.2 Le plan d'essais

La définition d'un plan d'essai accéléré dépend de plusieurs paramètres ([Hoang, 2003b], [Nelson, 1990a], [Vassilious and Mettas, 2001], [Meeker and Escobar, 1998], ...) :

- **Les stress d'accélération et les limites opérationnelles** : on appelle stress l'ensemble des conditions et facteurs susceptibles d'affecter le bon fonctionnement d'un produit. Les stress peuvent être de toute nature (mécanique, électronique, climatique, ...) et leurs durées de manifestation de tout ordre (constante, échelonnée, progressive, cyclique ou aléatoire). Les nombres et les niveaux des stress appliqués sont choisis en fonction du produit étudié avec la conservation du mécanisme de défaillance d'origine tout en accélérant suffisamment

l'essai. Les stress sont parfois désignés par les termes : contraintes et sollicitations. Les limites opérationnelles du produit sont déterminées par un essai aggravé par exemple (donnant les niveaux de stress extrêmes à ne pas dépasser afin d'éviter les fonctionnements dégradés des produits) ou changement de mécanisme de défaillance,

- **Les modes et mécanismes de défaillance** : dans un essai accéléré, les mécanismes d'endommagement provoqués doivent être représentatifs des conditions normales d'emploi. Chaque mode de défaillance peut être provoqué par un ou plusieurs types de contraintes,
- **Les nombres de dispositifs testés à chaque niveau de contrainte**, donnant la précision des estimations,
- **Un modèle de vie accélérée**, qui relie la durée de vie sous les conditions accélérées à celles sous les conditions normales d'utilisation, permet d'analyser les résultats d'essai pour estimer la fonction de fiabilité dans les conditions nominales.

La conception du plan d'essai (quels essais faut-il faire pour montrer que le système est fiable?) peut intervenir dès le début de la conception du système, dès que les fonctions requises du cahier des charges sont connues.

Les stress peuvent être appliqués sur les produits selon différents profils :

1. Le chargement constant (indépendant du temps) : on soumet chaque composant à un niveau de contrainte constant supérieur à la normale (figure 3.1),

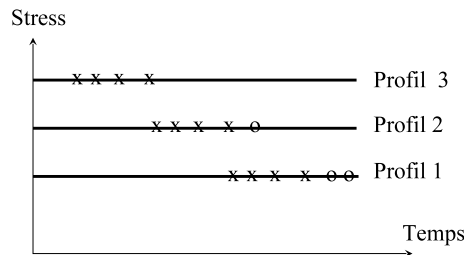


Figure 3.1 – Profil d'essai sous stress constant

2. Le chargement variable (en fonction du temps) : la contrainte choisie, évolue en fonction du temps. Elle peut être augmentée d'une manière échelonnée, progressive, cyclique ou aléatoire etc.:

- La contrainte échelonnée : la contrainte est appliquée de manière échelonnée par des niveaux croissants ou décroissants dans le temps (par paliers) jusqu'à l'apparition d'une défaillance (figure 3.2),

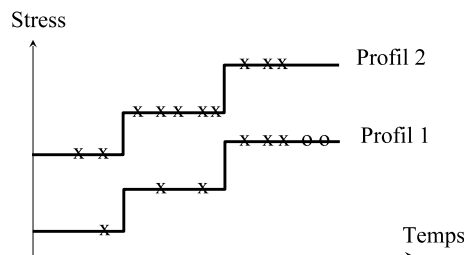


Figure 3.2 – Profil d'essai sous stress échelonné

- La contrainte progressive : la contrainte est augmentée de manière linéaire dans le temps (croissance linéaire),
- La contrainte cyclique : la contrainte est appliquée selon une amplitude et une fréquence données (figure 3.3),

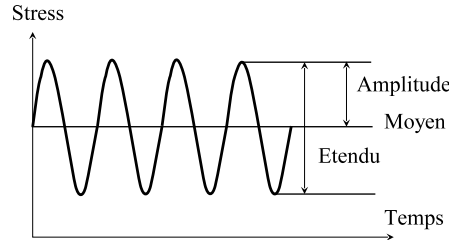


Figure 3.3 – Profil d'essai sous stress cyclique

- La contrainte aléatoire : la sollicitation suit un profil aléatoire (couramment utilisée en vibration)

Dans le cadre de cette thèse, nous allons utiliser le plan d'essais par chargement constant [Nelson, 1990a] avec des données censurées à droite, des niveaux de stress X_1, \dots, X_k constants accélérés : $X_1 < X_2 < \dots < X_k$, avec X_0 le stress usuel, ($X_0 < X_1$). On considère que k groupes d'unités sont observés. On teste le i -ème groupe sous le stress $X_i \in \varepsilon_0$ jusqu'à l'instant T_i . Le stress usuel X_0 n'est pas utilisé.

A chaque niveau de stress X_i on teste n_i produits et on note les temps de défaillance $t_{i1}, \dots, t_{iq_i}, q_i \leq n_i$, produits testés jusqu'à l'instant T_i ; il s'agit donc d'essais censurés à droite. Dans la figure 3.1, les temps de défaillance observés inférieurs à l'instant T_i sont marqués avec \mathbf{x} et les temps de défaillance possibles supérieurs à l'instant T_i avec \mathbf{o} .

Ce plan de test a les avantages suivants :

- simple car il sollicite les produits selon des stress constants,
- nombreux modèles disponibles pour traiter les résultats d'essai,
- plusieurs techniques d'estimation peuvent être utilisées.

Dans le cadre de cette thèse, nous proposons d'utiliser le modèle standard de vie accélérée qui a été très largement étudié par l'équipe de M. Nikulin ([Bagdonavicius and Nikulin, 1995b], [Bagdonavicius and Nikulin, 1995a], [Bagdonavicius and Nikulin, 1997b], [Bagdonavicius et al., 2000], [Bagdonavicius and Nikulin, 2001], [Bagdonavicius and Nikulin, 2001])

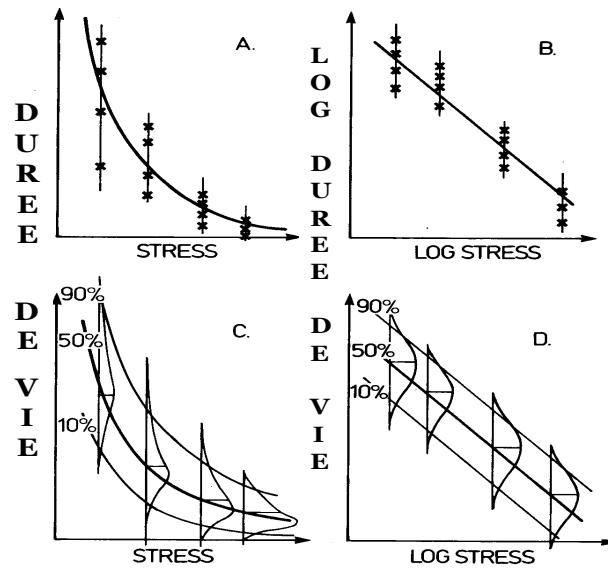


Figure 3.4 – Représentations graphiques d'un Modèle Statistique de vie Accéléré (durée de vie en fonction de la contrainte, tirée de [Nelson, 1990a])

3.2 Le modèle standard de vie accélérée (SVA)

Les modèles statistiques de vie accélérée sont généralement utilisés quand la relation exacte entre les contraintes appliquées et le temps de défaillance du composant est difficile à déterminer selon des principes physico-chimiques. Dans ce cas, les composants sont soumis à différents niveaux de contraintes et les instants de défaillance sont utilisés pour ajuster la distribution de vie la plus adéquate. Les instants de défaillance ont le même type de loi à chaque niveau de contrainte, en particulier dans les conditions normales de fonctionnement.

Les modèles statistiques de vie accélérée peuvent s'appliquer à plusieurs domaines comme celui du vivant (sciences médicales), de l'électronique et de la mécanique. Ce qui différencie les diverses applications sont les lois de fiabilité utilisées, les stress employés pour aggraver les essais et la nature des lois d'accélération.

Dans la littérature, on trouve plusieurs définitions théoriques des modèles Statistiques de Vie Accélérée ([Nelson, 1990a], [Bagdonavicius and Nikulin, 1995a], [Hoang, 2003b]). En général, ils sont constitués de deux composantes principales, la figure 3.4 illustre ce genre de modèles:

- **un modèle analytique "Durée de vie/stress"** appelé aussi loi d'accélération ou modèle d'accélération, traduisant la durée de vie "nominale" du produit soumis à l'essai en fonction des niveaux de contraintes appliquées. Cette durée de vie "nominale" est représentée par une caractéristique de la loi de fiabilité telle que la moyenne, la médiane, l'écart type, un quantile ou un quelconque paramètre de la loi.
- **une distribution statistique des durées de vie** : Dans un test accéléré, un modèle analytique seul, ne décrit pas le comportement des durées de vie du produit. À chaque niveau de contrainte (stress), le système a une distribution statistique de vie; d'où la combinaison : équation d'accélération et distribution de vie de base.

Devant une telle variété de modèles, un grand intérêt apparaît à unifier ces modèles dans un seul formalisme. C'est le but du modèle standard de vie accélérée que nous allons présenter dans

la sous-section suivante.

3.2.1 Principe

Supposons que la durée de vie $T_{X(\bullet)}$ sous n'importe quel stress $X(\bullet)$ est une variable aléatoire non négative absolument continue de fonction de survie:

$$R_{X(\bullet)}(t) = P\{T_{X(\bullet)} > t\}, t \geq 0 \quad (3.2)$$

Considérons un ensemble de stress ε . Formellement, on dit qu'un stress $X_1(\bullet)$ est supérieur à un stress $X_0(\bullet)$ si $R_{X_0(\bullet)}(t) > R_{X_1(\bullet)}(t)$ pour tout $t \geq 0$. Soit $\varepsilon_0 \subset \varepsilon$ un ensemble de stress constant dans le temps et $X_0 \in \varepsilon_0$. On peut définir la fonction inverse de $R_{X_0}(t)$ comme :

$$R_{X_0}^{-1}(p) = \inf\{t : R_{X_0}(t) \geq p\} \quad (3.3)$$

On note le produit de *convolution* entre $R_{X_0}^{-1}$ et $R_{X(\bullet)}$ comme $f_{X(\bullet)}(t) = R_{X_0}^{-1}R_{X(\bullet)}(t)$, avec $f_{X(\bullet)}(0) = 0$. Alors, pour tout $X(\bullet) \in \varepsilon$ on a :

$$P\{T_{X_0} \geq f_{X(\bullet)}(t)\} = P\{T_{X(\bullet)} \geq t\} \quad (3.4)$$

Il résulte que pour tout $X(\bullet) \in \varepsilon$ la probabilité de survivre jusqu'à l'instant t sous le stress $X(\bullet)$ est la même que la probabilité de survivre jusqu'à l'instant $f_{X(\bullet)}(t)$ sous le stress X_0 . On appelle la fonction $f_{X(\bullet)}(t)$ la *ressource utilisée sous le stress $X(\bullet)$ jusqu'à l'instant t* ou *fonction de transfert*.

Le *modèle standard de vie accélérée* (SVA) est défini sur ε s'il existe une fonction $r : \varepsilon \rightarrow \mathfrak{R}^+$ telle que pour tout $X(\bullet) \in \varepsilon$:

$$\frac{d}{dt}f_{X(\bullet)}(t) = r[X(t)] \quad (3.5)$$

La fonction r exprime une vitesse de dégradation. Le modèle SVA signifie que la vitesse d'utilisation de la ressource à l'instant t ne dépend que de la valeur du stress appliqué à l'instant t .

La formule 3.5 implique :

$$R_{X(\bullet)}(t) = R_{X_0}\left(\int_0^t r[X(\tau)] d\tau\right) \quad (3.6)$$

Dans le ce cas où $X(\tau) \equiv X = \text{constant}$, le modèle 3.6 implique :

$$R_X(t) = R_{X_0}(r(X)t) \quad (3.7)$$

Donc le stress ne change que l'échelle. Notons que $r(X_0) = 1$.

Le modèle SVA permet d'estimer la fiabilité d'un composant sous des contraintes expérimentalement indisponibles à partir de la fiabilité du même composant sous des conditions d'utilisation. L'analyse statistique du modèle SVA consiste à estimer les paramètres de ce modèle (fiabilité, taux de défaillance, ...). Les techniques d'estimation, dépendent de ce modèle statistique, c'est-à-dire de la connaissance de la fonction de survie R_{X_0} et du choix de la vitesse de dégradation r .

Si la fonction r est complètement inconnue, la fonction R_{X_0} ne peut pas être estimée même si l'on connaît la famille de distributions à laquelle R_{X_0} appartient.

Par exemple, si $R_{X_0}(t) = e^{[-(\frac{t}{\theta})^\alpha]}$ on obtiens $R_{X_i}(t) = e^{[-\left(\frac{r(X_i)}{\theta}t\right)^\alpha]}$. A partir de temps de défaillance on peut estimer le paramètres $\alpha, \frac{r(X_1)}{\theta}, \dots, \frac{r(X_k)}{\theta}$ mais puisque r est complètement inconnue, $r(X_0)$ et donc $S_{X_0}(t)$ ne peuvent pas être estimés.

Donc la fonction r doit être choisie dans une certaine classe de fonctions. Si le modèle 3.7 est vérifié sur un ensemble de stress ε , alors, pour tous X_i, X_j on a

$$R_{X_j}(t) = R_{X_i}\left(\frac{r(X_j)}{r(X_i)}t\right) = R_{X_i}\left(\rho(X_i, X_j)\right) \quad (3.8)$$

D'abord, supposons que le stress $x \in \mathfrak{R}$ soit unidimensionnel. Le taux de changement d'échelle est déterminé par la dérivée :

$$\delta(x) = \lim_{\Delta x \rightarrow 0} \frac{\rho(x, x + \Delta x) - \rho(x, x)}{\Delta x} = [\log r(x)]' \quad (3.9)$$

Donc pour tout $x \in \varepsilon$

$$r(x) = e^{\int_{x_0}^x \delta(v) dv} \quad (3.10)$$

Supposons que $\delta(x)$ soit proportionnelle à une fonction connue $u(x)$ du stress

$$\delta(x) = \beta_1 u(x), \beta_1 > 0 \quad (3.11)$$

Alors

$$r(x) = e^{\beta_0 + \beta_1 z(x)} \quad (3.12)$$

où $z(x)$ est une fonction primitive de $u(x)$, β_0, β_1 sont des paramètres inconnus.

Pour généraliser, si le stress est multidimensionnel : $X = (x_1, \dots, x_m)^T$, on considère des caractéristiques infinitésimales $\delta_i(x)$ données par :

$$\delta_i(X) = \lim_{\Delta x_i \rightarrow 0} \frac{\rho(X, X + \Delta x_i e_i) - \rho(X, X)}{\Delta x_i} = \frac{\partial \log r(X)}{\partial x_i} \quad (3.13)$$

où $e_i = \left(0, \dots, \underset{i}{1}, \dots, 0\right)^T$. L'unité est dans la i -ème coordonnée. Généralisant le cas unidimensionnel, $\delta_i(X)$ devient :

$$\delta_i(X) = \sum_{j=1}^{k_i} \beta_{ij} u_{ij}(X) \quad (3.14)$$

où les $u_{ij}(X)$ sont des fonctions connues et β_{ij} sont des constantes inconnues. Dans ce cas

$$r(X) = e^{\beta_0 + \sum_{i=1}^m \sum_{j=1}^{k_i} \beta_{ij} z_{ij}(X)} \quad (3.15)$$

avec $z_{ij}(X)$ des fonctions connues.

Alors, en généralisant, les modèles 3.6 et 3.7 peuvent s'écrire :

$$R_{X(\bullet)}(t) = R_{X_0} \left(\int_0^t e^{\beta^T Z(\tau)} d\tau \right) \quad (3.16)$$

où, pour des stress constants :

$$R_X(t) = R_{X_0} \left(t \cdot e^{\beta^T Z} \right) \quad (3.17)$$

où $\beta = (\beta_0, \dots, \beta_m)^T$ un vecteur de paramètres et $Z = (\varphi_0(X), \dots, \varphi_m(X))^T$ avec les fonctions φ_i spécifiées et la première composante z_0 égale à 1.

Dans la littérature ([Nelson, 1990a],...) on trouve une grande variété de modèles d'accélération, définis pour chaque type de composants et matériaux. Nous présentons quelques modèles d'accélération les plus utilisés pour des stress à profil constants, des cas particuliers de modèles SVA. Désormais, la durée de vie " nominale " est représentée par la moyenne.

3.2.2 Lois d'accélération courantes

3.2.2.1 Loi d'Arrhenius

Utilisée lorsque le mécanisme d'endommagement est sensible à la température (en particulier : diélectrique, semi-conducteur, batterie, lubrifiant et graisse, plastique et filament de lampe incandescente). La loi d'Arrhenius modélise la durée de vie τ du produit comme fonction de la température T :

$$\tau = A e^{\left(\frac{+B}{T}\right)} \quad (3.18)$$

avec :

- A et B constantes positives,
- T température absolue.

Le facteur d'accélération (FA) d'Arrhenius entre la durée de vie τ_1 pour une température T_1 et la durée de vie τ_2 pour une température T_2 est :

$$FA = \frac{\tau_1}{\tau_2} = e^{B \left(\frac{1}{T_1} - \frac{1}{T_2} \right)} \quad (3.19)$$

3.2.2.2 Loi de Peck

Utilisée lorsque le mécanisme d'endommagement est sensible à la température T et à l'humidité H (en particulier : composants électriques, conducteur aluminium et composants mécaniques soumis à la rupture). Le modèle est défini par :

$$\tau = A (H)^{-m} e^{\left(\frac{E_a}{kT}\right)} \quad (3.20)$$

avec

- T : température absolue en °K
- H : humidité en %
- k : $8.6171 \cdot 10^{-5}$ eV/K constante de Boltzmann
- E_a : énergie d'activation en eV
- m paramètre du modèle

3.2.2.3 Loi puissance inverse

Elle est utilisée lorsque le mécanisme d'endommagement est sensible à un stress particulier (par exemple : diélectrique, roulement à billes ou rouleaux, composants optoélectroniques, composants mécaniques soumis à la fatigue et filament de lampe incandescente).

La loi de puissance inverse décrit la cinétique d'une dégradation sous stress constant V , la durée de vie est donnée par l'équation :

$$\tau = \left(\frac{A}{V}\right)^\gamma \quad (3.21)$$

avec :

- V : stress constant,
- A et γ : constantes dépendantes de la défaillance et de l'essai.

Cas particulier : quand le stress est la température, la loi est appelée la loi de Coffin-Manson.

Le facteur d'accélération entre la durée de vie τ_1 pour un stress V_1 et la durée de vie τ_2 pour un stress V_2 est :

$$FA = \frac{\tau_1}{\tau_2} = \left(\frac{V_2}{V_1}\right)^\gamma \quad (3.22)$$

3.2.2.4 Loi d'Eyring

Utilisée lorsque le mécanisme d'endommagement est sensible à la température et un stress V (en particulier : composants électriques, conducteur aluminium et composants mécaniques soumis à la rupture). Le modèle est défini par :

$$\tau = \left(\frac{A}{T}\right) e^{\left(\frac{B}{kT}\right)} e^{V\left(C + \frac{D}{kT}\right)} \quad (3.23)$$

avec:

- T : température absolue
- V : Stress donné
- k : $8.6171 \cdot 10^{-5}$ eV/K constante de Boltzmann
- A, B, C et D les paramètres du modèle

Dans les paragraphes suivants, nous abordons l'estimation du modèle SVA selon l'approche paramétrique.

3.2.3 La construction d'un modèle paramétrique

Les modèles paramétriques sont généralement employés lorsqu'il existe une connaissance suffisante sur les lois de dégradation et les fonctions de fiabilité appropriées au produit à tester. Ainsi, le plan d'essai consiste à solliciter les composants à différents niveaux accélérés de stress. On fait l'hypothèse que la distribution des temps de défaillance est identique aux différents niveaux de stress (y compris dans les conditions initiales) et que le paramètre de forme de la distribution soit le même pour tous les niveaux de stress. Les temps de défaillance aux différents niveaux de stress sont utilisés pour déterminer la distribution de probabilité la plus appropriée [Shyur et al., 1999a].

En reprenant l'équation 2.19, on peut noter :

$R(t) = R(e^t)$, $t \in \mathfrak{R}$, $\sigma = \frac{1}{v}$, $v = (v_0, \dots, v_m)$, $\gamma_0 = \ln \eta - \beta_0$ et $\gamma_i = -\beta_i$, $i = (1, \dots, m)$, on peut écrire le modèle 3.16 sous la forme :

$$R_X(t) = R\left(\frac{\ln t - \gamma^T Z}{\sigma}\right) \quad (3.24)$$

Notons T_{ij} la durée de vie observée ou censurée de la j -ième unité du i -ième groupe, on a :

$T_{ij} = \ln(T_{ij} \wedge t_i)$, $\delta_{ij} = I\{T_{ij} \leq t_i\}$, $f(u) = -S'$, $\lambda(u) = \frac{f(u)}{S(u)}$ avec \wedge le minimum entre deux valeurs.

Dans ce cas, la fonction de vraisemblance est :

$$L(\gamma, \sigma | T) = \prod_{i=1}^k \prod_{j=1}^{n_i} \left[\frac{1}{\sigma} \lambda\left(\frac{T_{ij} - \gamma^T Z^{(i)}}{\sigma}\right) \right]^{\delta_{ij}} R\left(\frac{T_{ij} - \gamma^T Z^{(i)}}{\sigma}\right) \quad (3.25)$$

En dérivant par rapport à γ_i et σ la fonction $\ln L(\gamma, \sigma)$, on obtient :

$$\begin{aligned} U_l(\gamma, \sigma) &= \frac{\partial \ln L(\gamma, \sigma)}{\partial \gamma_l} = \frac{1}{\sigma} \sum_{i=1}^k z_{il} \sum_{j=1}^{n_i} a_{ij}(\gamma, \sigma), \quad (l = 1, \dots, m), \\ U_{m+1}(\gamma, \sigma) &= \frac{\partial \ln L(\gamma, \sigma)}{\partial \sigma} = \frac{1}{\sigma} \sum_{i=1}^k \sum_{j=1}^{n_i} \{v_{ij}(\gamma, \sigma) a_{ij}(\gamma, \sigma) - \delta_{ij}\}, \end{aligned} \quad (3.26)$$

avec

$$v_{ij}(\gamma, \sigma) = \frac{T_{ij} - \gamma^T Z^{(i)}}{\sigma}, \quad a_{ij}(\gamma, \sigma) = \lambda(v_{ij}(\gamma, \sigma)) - \delta_{ij} (\ln \lambda)'(v_{ij}(\gamma, \sigma))$$

Les estimateurs du maximum de vraisemblance $\hat{\sigma}, \hat{\gamma}$ peuvent être obtenus en résolvant le système d'équations :

$$U_l(\gamma, \sigma) = 0 \quad (l = 1, \dots, m+1) \quad (3.27)$$

et l'estimateur de la fonction de survie R_{X_0} est

$$\hat{R}_{X_0}(t) = R\left(\frac{\ln t - \hat{\gamma}^T Z^{(0)}}{\hat{\sigma}}\right) \quad (3.28)$$

Pour calculer l'intervalle de confiance $(1-\alpha)$ sur l'estimateur de la fonction de fiabilité, on définit la matrice d'information de Fisher $I^L(\gamma, \sigma)$ avec les éléments suivants :

$$\begin{aligned}
 I_{ls}^L(\gamma, \sigma) &= -\frac{\partial^2 \ln L(\gamma, \sigma)}{\partial \gamma_l \partial \gamma_s} \\
 &= \frac{1}{\sigma^2} \sum_{i=1}^k z_{il} z_{is} \sum_{j=1}^{n_i} c_{ij}(\gamma, \sigma), \quad l, s = 0, \dots, m
 \end{aligned} \tag{3.29}$$

$$\begin{aligned}
 I_{l,m+1}^L(\gamma, \sigma) &= -\frac{\partial^2 \ln L(\gamma, \sigma)}{\partial \gamma_l \partial \sigma} \\
 &= \frac{1}{\sigma} U_l(\gamma, \sigma) + \frac{1}{\sigma^2} \sum_{i=1}^k z_{il} \sum_{j=1}^{n_i} v_{ij}(\gamma, \sigma) c_{ij}(\gamma, \sigma), \quad l = 0, \dots, m
 \end{aligned} \tag{3.30}$$

$$\begin{aligned}
 I_{m+1,m+1}^L(\gamma, \sigma) &= -\frac{\partial^2 \ln L(\gamma, \sigma)}{\partial \sigma^2} \\
 &= \frac{2}{\sigma} U_{m+1}(\gamma, \sigma) + \frac{1}{\sigma^2} \sum_{i=1}^k \sum_{j=1}^{n_i} \left(v_{ij}^2(\gamma, \sigma) c_{ij}(\gamma, \sigma) + \delta_{ij} \right)
 \end{aligned} \tag{3.31}$$

avec : $c_{ij}(\gamma, \sigma) = \lambda'(v_{ij}(\gamma, \sigma)) - \delta_{ij}(\ln \lambda)''(v_{ij}(\gamma, \sigma))$.

On observe que l'estimateur \hat{R} est une fonction régulière de γ et σ . On définit les fonctions :

$$Q_{X_0}(t) = \ln \frac{R_{X_0}(t)}{1 - R_{X_0}(t)} \quad \text{et} \quad \hat{Q}_{X_0}(t) = \ln \frac{\hat{R}_{X_0}(t)}{1 - \hat{R}_{X_0}(t)} \tag{3.32}$$

La fonction Q_{X_0} prend ses valeurs dans \mathfrak{R} donc la vitesse de convergence de \hat{Q}_{X_0} vers sa loi limite est plus grande que celle de \hat{R}_{X_0} vers sa loi limite. La loi de

$$\frac{\hat{Q}_{X_0}(t) - Q_{X_0}(t)}{\hat{\sigma}_{Q_0}}$$

est approximée par la loi normale $N(0, 1)$, avec :

$$\begin{aligned}
 (\hat{\sigma}_{Q_0})^2 &= \left(\frac{\partial \hat{Q}_{X_0}}{\partial \hat{\gamma}_0}, \dots, \frac{\partial \hat{Q}_{X_0}}{\partial \hat{\gamma}_m}, \frac{\partial \hat{Q}_{X_0}}{\partial \hat{\sigma}} \right) \left(I^L(\gamma, \sigma)^{-1} \right) \left(\frac{\partial \hat{Q}_{X_0}}{\partial \hat{\gamma}_0}, \dots, \frac{\partial \hat{Q}_{X_0}}{\partial \hat{\gamma}_m}, \frac{\partial \hat{Q}_{X_0}}{\partial \hat{\sigma}} \right) \\
 &= \left(\frac{R'(R^{-1}(\hat{R}_{X_0}(t)))}{\hat{\sigma}^2 \hat{R}_{X_0}(t) (1 - \hat{R}_{X_0}(t))} \right)^2 \times \left[\hat{\sigma}^4 \sum_{l=0}^m \sum_{s=0}^m z_{0l} z_{0s} I_L^{ls}(\hat{\gamma}, \hat{\sigma}) \right. \\
 &\quad \left. - 2\hat{\sigma}^2 \hat{\gamma}^T Z^{(0)} \sum_{l=0}^m z_{0l} I_L^{l,m+1}(\hat{\gamma}, \hat{\sigma}) + (\hat{\gamma}^T Z^{(0)})^2 I_L^{m+1,m+1}(\hat{\gamma}, \hat{\sigma}) \right]
 \end{aligned} \tag{3.33}$$

avec $I_L^{l,s}$ les termes de la matrice $(I^L(\gamma, \sigma))^{-1}$.

L'intervalle de confiance bilatéral symétrique approximatif de niveau de confiance $(1-\alpha)$ pour $\hat{R}_{X_0}(t)$ est :

$$\left(1 + \frac{1 - \hat{R}_{X_0}(t)}{\hat{R}_{X_0}(t)} \exp \left\{ \pm \hat{\sigma}_{Q_0} w_{1-\alpha/2} \right\} \right)^{-1} \tag{3.34}$$

avec w_α la α -quantile de la loi $N(0, 1)$.

Alors, à partir du modèle générique 3.7 on peut créer une multitude de modèles dérivés en choisissant différentes lois d'accélération Arrhenius, Eyring, Peck, Le plus souvent, ces modèles sont appliqués de la façon suivante : on suppose que la fonction $R_{X_0}(t)$ appartient à une certaine classe Γ de répartitions et on observe plusieurs groupes des composants sous des stress accélérés en conformité avec le plan d'essais classique présenté dans 3.1.2. On estime alors les paramètres inconnus $\hat{\gamma}$ et $\hat{\sigma}$ de $r(X)$ et la fonction $R_{X_0}(t)$. Un estimateur de la fonction de fiabilité sous le stress "normal" X_0 est obtenu en changeant dans 3.7 X par X_0 ainsi que les estimateurs des paramètres inconnus.

Les points faibles de cette méthodologie sont les suivants [Bagdonavicius and Nikulin, 1997a]:

1. impossibilité de vérifier l'hypothèse que la loi obtenue en conditions sévères est celle observée dans le cadre de stress "normal" à partir de telles expériences,
2. choix limité de la fonction $r(X)$ (log-linéaire en particulier),
3. erreur de prédiction du modèle de régression car le stress nominal X_0 n'est pas utilisé pendant les expériences.

Pour traiter des deuxième et troisième points, des plans ont été proposés, basés sur le modèle non-paramétrique par exemple, où quasiment aucune hypothèse n'est faite sur la classe d'appartenance de $r(X)$ [Tebbi, 2005] et de la fonction de fiabilité. Ces plans ne seront pas traités dans le cadre de cette thèse car ils ne présentent pas d'intérêt dans le contexte abordé. En effet, nous nous plaçons dans le cas d'essais de qualification permettant d'utiliser les modèles SVA complets (fonction de fiabilité et loi d'accélération connues) pour estimer la fonction de fiabilité d'un produit soumis à un environnement complexe.

3.3 Application numérique sur le modèle de durée de vie des roulements à billes

Dans cette section, les essais accélérés sont appliqués pour évaluer la fiabilité de roulements à billes qui ont fait l'objet de nombreux études et développement. Ainsi, il est possible aisément de développer un simulateur numérique d'essai permettant d'obtenir des instants de défaillance dans différentes conditions d'essai.

Pour commencer nous présentons les modèles de fiabilité des roulements. Ensuite, nous définissons le modèle de simulation et détaillons la procédure d'estimation du modèle SVA paramétrique à partir des résultats d'essai simulés.

Les roulements à billes sont couramment utilisés dans le guidage en rotation de nombreux équipements et machines (moteurs, compresseurs, réducteurs, ...). Ces composants ont fait l'objet de nombreuses études et analyses ayant permis de bâtir des modèles de fiabilité. Ainsi, il est courant de caractériser la fiabilité des roulements par la durée L_{10} , qui est également notée B_{10} . Cette durée représente le temps atteint sans défaillance par au moins 90% d'une population homogène de roulements dans des conditions d'utilisation identiques. En pratique, la durée L_{10} (en millions de cycles) est caractérisée par le modèle de Palmgreen :

$$L_{10} = \left(\frac{C}{P} \right)^p \quad (3.35)$$

où

- C est la charge dynamique de base,
- P est la charge radiale équivalente,
- $p = 3$ pour les roulements à billes et $p = 10/3$ pour les roulements à rouleaux.

La fonction de fiabilité est usuellement définie par une loi de Weibull combinée au modèle de Palmgreen :

$$R(u, P) = e^{-0.105 \left(\frac{u}{L_{10}}\right)^{1.5}} \tag{3.36}$$

Le paramètre d'échelle η est donné par :

$$\eta(P) = L_{10} \left(\frac{1}{0.105}\right)^{\frac{1}{1.5}} \tag{3.37}$$

ou η_H en heures:

$$\eta_H(P) = L_{10} \frac{10^6}{60.N} \left(\frac{1}{0.105}\right)^{\frac{1}{1.5}} \tag{3.38}$$

Avec N la vitesse de rotation. Le paramètre de forme β est fixé à 1.5.

3.3.0.1 Définition du modèle de simulation

Le simulateur de résultats d'essai est défini en considérant les caractéristiques d'un roulement à billes données dans le tableau 3.1.

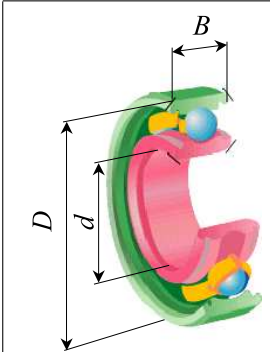
| | |
|---|--|
|  | $D = 32$ mm : diamètre extérieur |
| | $d = 15$ mm : diamètre intérieur |
| | $B = 8$ mm : largeur |
| | $C = 400$ daN : charge dynamique de base |
| | $N = 1500$ tr/min : vitesse de rotation |

Table 3.1 – Caractéristiques du roulement

Pour illustrer le modèle SVA, la charge P est utilisée pour sévérer les conditions d'essai. Le plan d'essai utilisé est à deux niveaux constants de charge P_1 et P_2 , sans censure et des tailles d'échantillon identiques aux deux niveaux $n_1=n_2=20$. Ainsi, les paramètres de simulation sont définis dans le tableau 3.2.

| | Charge P en daN | Paramètre d'échelle η (eq 3.38) en heures | Paramètre de forme β |
|----------------------------|-------------------|---|----------------------------|
| conditions opérationnelles | 125 | 1635.86 | 1.5 |
| Test 1 | 150 | 946.68 | 1.5 |
| Test 2 | 175 | 596.16 | 1.5 |

Table 3.2 – Paramètres de Simulation

De ces paramètres de simulation, les valeurs N_{ij} sont simulées des lois de Weibull suivantes :

$$\text{Test 1: } N_{1j} \propto W(946.68, 1.5) \quad \text{Test 2: } N_{2j} \propto W(596.16, 1.5)$$

Les résultats de simulation sont donnés dans les tableaux 3.3.

| | | | | | | | | | | |
|--------|-------|-------|-------|-------|--------|--------|--------|--------|--------|--------|
| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
| Test 1 | 156.5 | 280.4 | 378.3 | 421.5 | 433.2 | 558.4 | 601.8 | 699.1 | 716.4 | 726.9 |
| Test 2 | 499.5 | 534.9 | 794 | 808.4 | 825.4 | 834.3 | 868.8 | 1126.2 | 1143.5 | 1492.2 |
| | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 | 20 |
| Test 1 | 749.6 | 778.8 | 800.9 | 884.9 | 1061.5 | 1229.9 | 1269.3 | 1468.9 | 1804.2 | 1910.1 |
| Test 2 | 499.5 | 534.9 | 794 | 808.4 | 825.4 | 834.3 | 868.8 | 1126.2 | 1143.5 | 1492.2 |

Table 3.3 – Résultats de simulations (N en heures)

Le modèle SVA utilisé pour l'analyse des résultats d'essai est défini par :

$$R_P(u) = e^{-\left(\frac{u}{e^{\beta_0 + \beta_1 \ln(P)}}\right)^\nu}$$

Les estimateurs ponctuels du modèle SVA sont obtenus par la méthode du maximum de vraisemblance.

| | | |
|------------------------------|------------------------------|--------------------------|
| $\hat{\beta}_0$ | $\hat{\beta}_1$ | $\hat{\nu}$ |
| 18.2422 | -2.2758 | 1.7086 |
| $\beta_0^{\text{théorique}}$ | $\beta_1^{\text{théorique}}$ | $\nu^{\text{théorique}}$ |
| 21.8849 | -3 | 1.5 |

Table 3.4 – Estimateurs ponctuels

Pour finir, la fonction de fiabilité dans les conditions opérationnelles ($P_0=125$ daN) peut être estimée (voir 3.5).

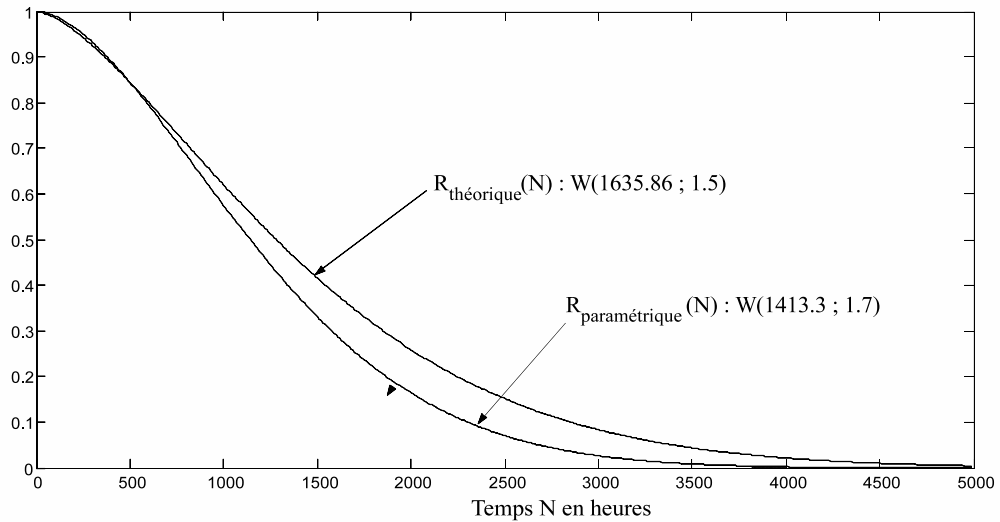


Figure 3.5 – Fonction de fiabilité dans les conditions opérationnelles ($P_0=125$ DaN)

Pour analyser le comportement asymptotique des estimateurs, une simulation de Monte Carlo est répétée (20 simulations avec des tailles d'échantillons $n_1 = n_2 = 100$ unités). Les estimateurs de paramètres inconnus sont caractérisés par leur moyenne et écart-type (Tableau 3.5).

| paramètre | moyenne | écart-type |
|-----------------|---------|------------|
| $\hat{\beta}_0$ | 20.98 | 2.96 |
| $\hat{\beta}_1$ | -2.82 | 0.59 |
| $\hat{\nu}$ | 1.49 | 0.099 |
| $\hat{\eta}$ | 1574.3 | 216.7 |

Table 3.5 – Résultats de simulation de Monte Carlo

La figure 3.6 représente les fonctions de fiabilité $R_{P_0}(N)$ dans les conditions opérationnelles estimées à chaque répétition de la simulation de Monte Carlo.

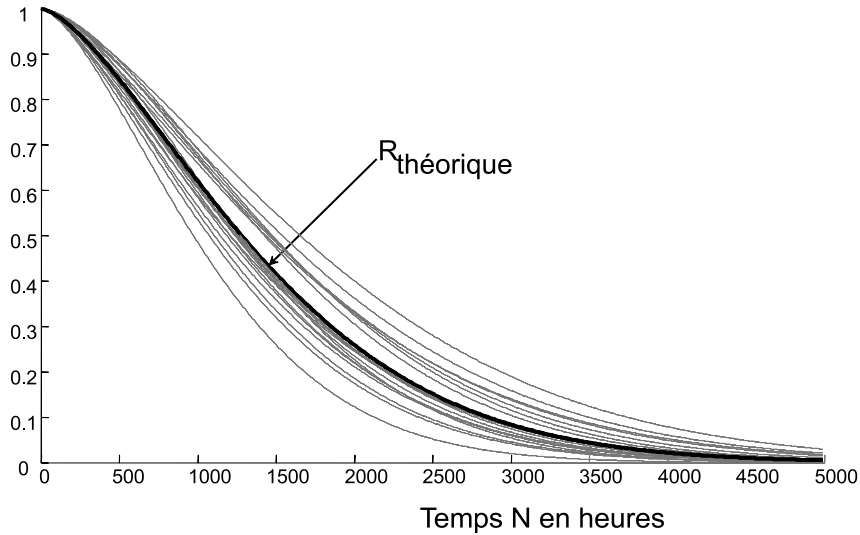


Figure 3.6 – Fonctions de fiabilité pour 20 répétitions de la simulation de monte Carlo

On observe que les fonctions de fiabilité estimées encadrent correctement la fonction de fiabilité théorique avec une variabilité modérée malgré les tailles d'échantillons importantes $n_1 = n_2 = 100$ unités.

Dans la simulation précédente, tous les roulements étaient défaillants. Dans la simulation suivante (avec les mêmes paramètres de simulation et 200 répétitions), l'estimation des paramètres de modèle paramétrique en présence de données censurées ($T_{censure} = 900$ h et 300 h) et pour différentes tailles d'échantillon ($n=10, 15, 25$ et 50) est étudiée. Pour simplifier l'analyse, on considère l'estimation de la probabilité de défaillance P_f , tirée du modèle paramétrique SVA, pour une durée de fonctionnement de $t=100$ h et dans les conditions opérationnelles ($P_0=125$ daN):

$$P_f = 1 - e^{-\left(\frac{t}{e^{\beta_0 + \beta_1 \cdot \ln(P)}}\right)^\beta}$$

Les figures 3.7 et 3.8 donnent l'évolution de l'estimation de P_f (estimateur ponctuel et intervalle de confiance à 90%) en fonction de la taille de l'échantillon et le temps de censure (valeur théorique en trait plein).

On peut remarquer que l'estimation P_f est sensible à de faibles tailles d'échantillon et aux temps de censure courts. On montre évidemment qu'avec l'augmentation de la taille d'échantillon les estimateurs ponctuels convergent vers les valeurs théoriques, utilisées pour la simulation, et que les intervalles de confiance (au niveau de confiance 90%) diminuent rapidement.

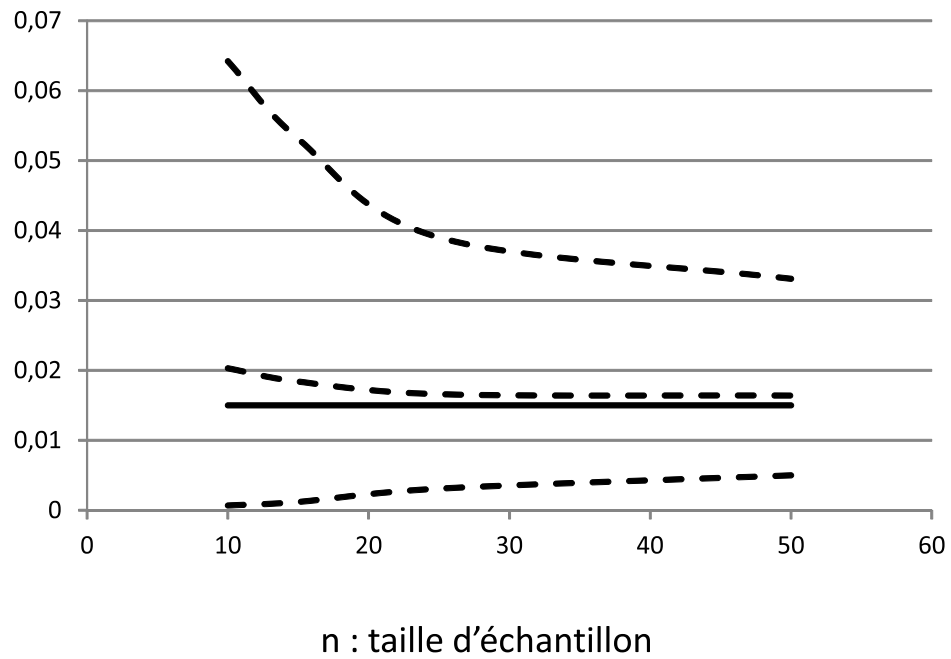


Figure 3.7 – Estimation de Pf au niveau de confiance 90% pour une censure à 900 h

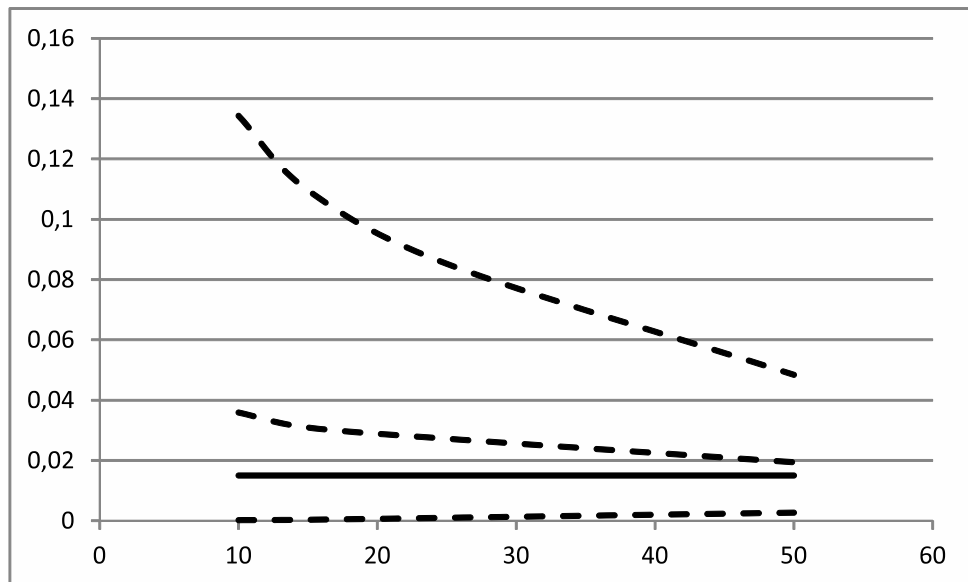


Figure 3.8 – Estimation de Pf au niveau de confiance 90% pour une censure à 300 h

3.4 Conclusion

Ce chapitre a présenté les essais accélérés comme une alternative pour d'accéder à des estimations, de bonne qualité et rapidement, de la fiabilité lorsque les produits à tester sont très fiables conduisant à des censures importantes s'ils sont testés dans les conditions nominales. Une des plus importantes contraintes de ces essais est de ne pas modifier le mode de défaillance par la sévérisation des conditions d'essai. Le plan étudié est celui par régression qui a pour effet de propager les incertitudes d'estimation obtenues aux niveaux accélérés à celui nominal. Ainsi la figure 3.9 illustre le phénomène de propagation d'incertitudes en considérant des couples d'échantillons différents.

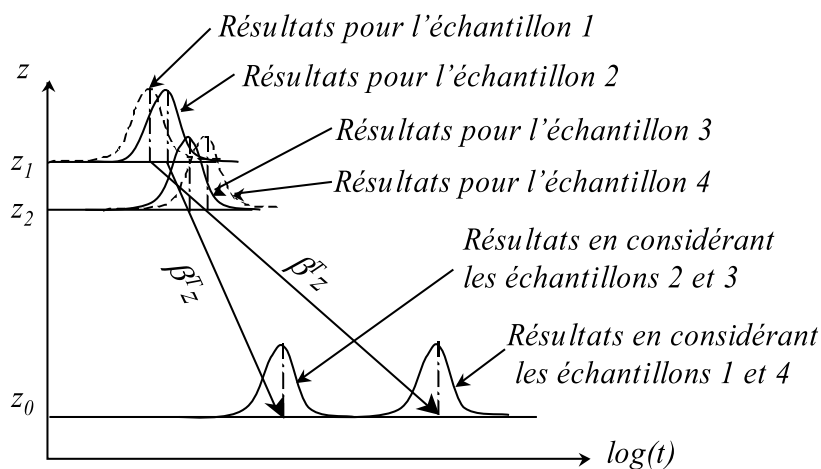


Figure 3.9 – Propagation de l'incertitude

Pour maîtriser les incertitudes, il est possible de rechercher la meilleure configuration de plan d'essai en minimisant l'incertitude d'estimation. De nombreux auteurs traitent la problématique de l'optimisation des plans d'essais accélérés [O'Connor, 2003b], [Guerin et al., 2001], [Arizono et al., 2008].

Une autre possibilité pour augmenter la qualité d'estimation sur les paramètres de modèles est d'introduire toute la connaissance disponible sur la fiabilité du système à étudier. En effet, lors du développement d'un produit de nombreuses études sont réalisées afin d'évaluer sa fiabilité (estimation prévisionnelle de la fiabilité par des standards de fiabilité (RdF2000, Mil Hdbk 217, Fides), AMDEC, ...). De plus, assez souvent un nouveau produit est une amélioration d'un ancien se trouvant en exploitation depuis longtemps et pour lequel nous avons des informations sur sa tenue en terme de fiabilité. D'autre part, il est toujours possible d'interroger un ou plusieurs experts sur le niveau de fiabilité potentiel du produit.

Aussi, nous proposons d'utiliser une démarche bayésienne, pour estimer les modèles standards de vie accélérée, permettant d'intégrer l'information disponible sur la fiabilité du produit. L'objectif est d'obtenir des estimations de bonne qualité (i.e un intervalle de confiance "raisonnable") pour des tailles d'échantillon faibles. Le chapitre suivant présente la formalisation de cette démarche.

Estimation bayésienne

4.1 Introduction

La *connaissance* est l'état de celui qui connaît ou sait quelque chose. Ce qu'est la connaissance fait encore débat chez les philosophes : il n'y a pas de définition unique sur laquelle ils s'accordent.

On distingue traditionnellement trois types de connaissance :

1. La connaissance *propositionnelle* est le fait de savoir qu'une certaine proposition est vraie, par exemple, savoir que la Terre est ronde,
2. La connaissance *objectuelle*, aussi appelée *acquaintance*, est le fait de connaître une chose, par exemple, connaître Paris,
3. Le *savoir faire* est le fait d'être capable de réussir une action, par exemple, savoir faire des crêpes.

La définition de la *connaissance propositionnelle* est celle qui a le plus attiré l'attention des philosophes. Ils s'accordent généralement sur le fait qu'une connaissance est une croyance qui est vraie, mais aussi qu'elle n'est pas seulement cela. Une autre condition est requise: il faut qu'elle soit justifiée, ou non accidentelle, ou fiable, ou certaine. C'est sur cette condition supplémentaire que les philosophes sont en désaccord.

Le terme *a priori* désigne les connaissances supposées par l'expérience, et qui ne peuvent se réduire à elles. Le concept d'*a priori* au sens kantien désigne ce qui est pensé comme nécessaire et universel et qui, à ce titre, peut être conçu "*indépendamment/indépendant de l'expérience*". Une connaissance est donc *a priori* si elle peut être prouvée sans aucune référence à l'expérience.

Comme nous l'avons montré précédemment, les essais d'estimation sont coûteux en temps. Ainsi, il est indispensable d'optimiser les campagnes d'essai en sachant que les niveaux de performances souhaitées à valider sont élevés (probabilité de défaillances et taux de défaillance très faibles) et que la taille des échantillons est telle que l'on a peu de chance d'observer des défaillances.

Dans cette optique, l'approche Bayésienne s'impose puisqu'elle va permettre :

1. d'utiliser le maximum de connaissances disponibles en exploitant les résultats des essais sur des produits similaires et les données issues de l'expertise,
2. d'obtenir des intervalles de confiance plus resserrés par rapport au cas classique,
3. d'intégrer à la connaissance les résultats des séries d'essais au fur et à mesure du développement du produit,
4. d'exploiter le cas du "zéro" défaillance difficile à traiter avec une approche statistique classique qui demanderait une taille d'échantillons très importante.

Il existe de nombreux travaux traitant les techniques bayésiennes. Nous pouvons citer les travaux fondamentaux : [Robert, 1992], [Robert, 2001], [Congdon, 2001], [Ibrahim et al., 2001]... et les travaux concernant les applications à la fiabilité : [Ringler, 1979b], [Procaccia et al., 1992b], [Procaccia and Morilhat, 1996b], [Lannoy and Procaccia, 1994b], [Lannoy and Procaccia, 2001], [Sander and Badoux, 1991b], [Marts and Walter, 1982], [Ross, 2003].... Le problème difficile à résoudre dans le cadre des techniques bayésiennes reste la modélisation de la connaissance *a priori* (connaissance initialement disponible). Pourtant, il existe de nombreuses sources de données de fiabilité (des études de sûreté de fonctionnement, étude d'évaluation de risques, d'avis d'expert, du retour d'expérience et des recueils de données).

L'inférence Bayésienne est fondée sur la manipulation d'énoncés probabilistes. Ces énoncés doivent être clairs et concis afin d'éviter toute confusion. Les méthodes bayésiennes se distinguent des méthodes dites standards par l'application systématique de règles formelles de transformation des probabilités.

On nomme inférence Bayésienne la démarche logique permettant de calculer ou réviser la probabilité d'une hypothèse. Cette démarche est régie par l'utilisation de règles strictes de combinaison des probabilités, desquelles dérive le théorème de Bayes. Dans la perspective Bayésienne, une probabilité n'est pas interprétée comme le passage à la limite d'une fréquence, mais plutôt comme la traduction numérique d'un état de connaissance (le degré de confiance accordé à une hypothèse).

Le théorème de Bayes est une conséquence simple des définitions des probabilités conditionnelles. La forme continue de théorème Bayes pour la variable aléatoire θ sur le domaine Ω avec t_i , ($i = 1..n$), des résultats d'essai, est :

$$\pi_{apo}(\theta/t_i) = \frac{L(t_1, \dots, t_n/\theta) \pi(\theta)}{\int_{\Omega} L(t_1, \dots, t_n/\theta) \pi(\theta) d\theta} \quad (4.1)$$

avec comme notations $L(t_1, \dots, t_n/\theta)$ la vraisemblance, $\pi(\theta)$ la fonction qui modélise l'information a priori et $\pi_{apo}(\theta/t_i)$ la fonction a posteriori.

Dans l'approche Bayésienne, le paramètre θ à estimer est une variable aléatoire sur l'espace probabilisable Ω et admet une densité de probabilité.

L'incertitude sur l'occurrence d'un événement est modélisée par une probabilité subjective, déterminée à partir de connaissances sur des composants similaires ou d'avis d'experts : c'est l'information a priori $\pi(\theta)$.

La fonction de vraisemblance va permettre de valoriser les données issues du retour d'expérience ou d'essais sur le composant étudié : c'est l'information objective $L(t_1, \dots, t_n/\theta)$.

Le théorème de Bayes (équation 4.1) va permettre de combiner l'information subjective avec l'information objective dans une information a posteriori $\pi_{apo}(\theta/t_i)$.

La difficulté d'utiliser les essais accélérés, au cours de phase de conception, est due à la faible taille de l'échantillon à tester. Dans ce contexte, l'approche Bayésienne peut être utilisée pour intégrer dans le processus d'estimation toutes les connaissances disponibles sur l'accélération du modèle de vie (le taux de défaillance, l'énergie d'activation, ...). L'application concrète du théorème de Bayes dépend fortement de la capacité à construire une distribution a priori.

Des nombreux articles traitent l'inférence Bayésienne sur les paramètres du modèle de vie accélérée. Dans [Bris, 2000], le cas des essais de démonstration de fiabilité est étudié en considérant l'hypothèse de mortalité exponentielle intégrant un facteur d'accélération. Dans [Madi, 1997], l'estimation de paramètres du modèle de vie accélérée est obtenue en utilisant l'inférence Bayésienne et l'échantillonnage de Gibbs. Dans [Somerville et al., 1997], les auteurs étudient les essais accélérés définis par une conception des expériences en utilisant la distribution a priori de Dirichlet. Dans [Dorp and Mazzuchi, 2004], l'approche Bayésienne est appliquée aux tests accélérés dans la définition des connaissances a priori directement aux niveaux accélérés. Dans [Dorp et al., 1996], les auteurs étudient l'inférence Bayésienne pour les essais accélérés effectués en contraintes échelonnées.

L'ensemble des travaux concerne des applications et études particulières de l'inférence bayésienne. Dans la suite de ce chapitre, nous proposons d'étudier l'estimation bayésienne des paramètres inconnus des modèles SVA. Nous nous intéresserons tout particulièrement à l'application

aux plans par régression en considérant les modèles paramétrique et semi-paramétrique. L'extension aux autres plans et modèles est quasi immédiate car le formalisme utilisé est commun. Une étude sur l'estimation est réalisée avec une comparaison entre les estimateurs classiques et MAP (Maximum of A Priori).

La vraie difficulté, lorsque l'on souhaite mettre en place une démarche bayésienne, est la construction de la loi a priori qui synthétise toute la connaissance disponible sur les paramètres du modèle. Aussi, nous proposons une démarche permettant la construction de celle-ci. Pour finir, nous étudierons la problématique de l'incompatibilité entre la connaissance a priori sur la fiabilité du produit et les premiers résultats d'essais. Aussi, nous proposons un facteur de compatibilité qui indique le degré de cohérence entre l'a priori et la vraisemblance. Cet indicateur est ensuite utilisé, lorsque l'incohérence est révélée, pour pondérer l'a priori afin de diminuer son informativité.

4.2 L'inférence Bayésienne dans le cadre de modèles SVA

L'intérêt du théorème de Bayes (équation 4.1) dans le cadre d'essais est de compenser la faible taille d'échantillon à tester et les censures nombreuses par l'intégration de l'information préalable obtenue par l'expertise ou la prise en compte de résultats d'essais des produits identiques, semblables ou de premières générations. L'information préalable en condition normale d'utilisation du produit sera assimilée à un résultat d'essais hypothétique (a priori) combiné aux valeurs de tests réels en utilisant l'inférence bayésienne.

On peut dire que la distribution a posteriori $\pi_{apo}(\theta/t_i)$ est le produit entre l'information priori $\pi(\theta)$ et l'information obtenue à partir des essais, modulés par la fonction de vraisemblance $L(t_1, \dots, t_n/\theta)$, le tout divisé par un facteur de normalisation afin d'obtenir une fonction densité de probabilité.

L'application du théorème de Bayes à l'estimation de la fiabilité d'un système consiste à appliquer l'inférence Bayésienne représentée dans la figure 4.1.

La figure 4.2 montre un exemple de représentation de ces différentes distributions.

En inférence Bayésienne, le choix crucial est celui de la distribution a priori. Lorsqu'on souhaite dissocier l'apport des données de celui des croyances initiales, on peut prendre une distribution qui ne reflète pas ces croyances, mais qui au contraire formalise un état d'ignorance (distribution neutre ou "non informative"); la distribution finale exprime alors l'apport propre des données: on obtient ainsi, dans la tradition de Laplace à Jeffreys, l'Analyse Bayésienne des Données. Un résultat fondamental est la ré-interprétation laplacienne du seuil observé, pour les modèles d'échantillonnage usuel.

Nous rappelons que, lorsque le nombre d'observations devient important, la subjectivité de la connaissance a priori perd de son importance au profit des données observées. Par contre, dans le cas d'essais où aucune défaillance n'est observée, l'inférence Bayésienne apporte une solution [Procaccia et al., 1992b], [Lannoy and Procaccia, 1994b], [Sander and Badoux, 1991b], [Usureau et al., 2004].

Dans le cadre de cette thèse, nous proposons d'étudier l'estimation Bayésienne des paramètres des modèles Standard de Vie Accélérée (SVA) paramétriques.

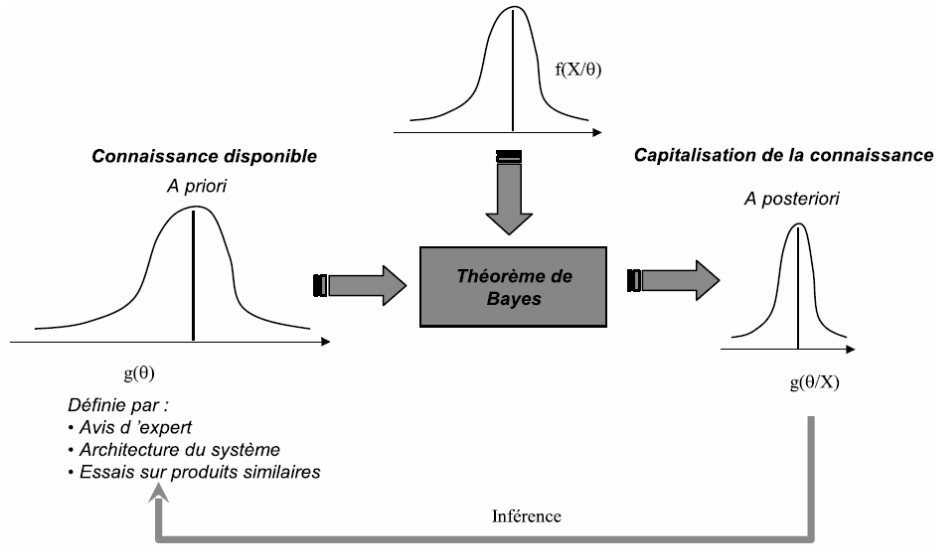


Figure 4.1 – Inférence Bayésienne

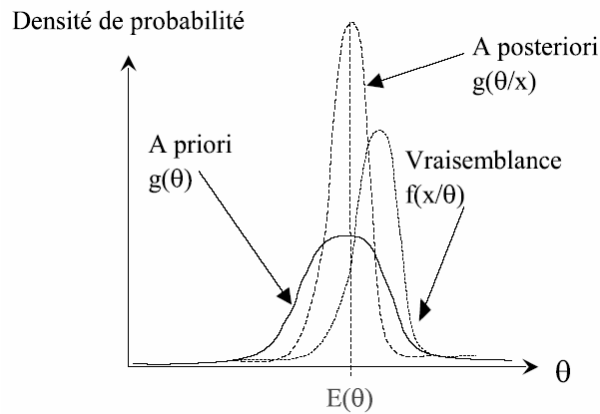


Figure 4.2 – Représentation des distributions a priori, vraisemblance et a posteriori

4.2.1 Application de l'inférence Bayésienne au modèle SVA paramétrique

4.2.1.1 Formulation du problème

L'intégration de l'inférence Bayésienne dans le modèle SVA paramétrique présenté dans la section 3.2.3 nous permet de compléter le cadre d'analyse de données de fiabilité proposé par [Nikulin et al., 2007].

La vraisemblance, donnée par l'équation 3.25, est :

$$L(\gamma, \sigma | T) = \prod_{i=1}^k \prod_{j=1}^{n_i} \left[\frac{1}{\sigma} \lambda \left(\frac{T_{ij} - \gamma^T Z^{(i)}}{\sigma} \right) \right]^{\delta_{ij}} R \left(\frac{T_{ij} - \gamma^T Z^{(i)}}{\sigma} \right) \quad (4.2)$$

Les paramètres inconnus à estimer sont σ et le vecteur γ de m éléments, considérés comme

des variables aléatoires. On considère une information a priori sur les paramètres inconnus modélisée par les fonctions de densité $\pi(\gamma_i)$ et $\pi(\sigma)$. Dans l'hypothèse des variables aléatoires indépendantes, on peut caractériser la totalité de l'information a priori sur les paramètres par :

$$\pi(\gamma, \sigma) = \pi(\sigma) \prod_{i=1}^m \pi(\gamma_i) \quad (4.3)$$

En intégrant l'équation de la fonction de vraisemblance définie par l'équation 3.25 dans la forme générale continue du théorème du Bayes (équation 4.1), on obtient :

$$\pi_{apo}(\gamma, \sigma | T) = \frac{\left[\prod_{i=1}^k \prod_{j=1}^{n_i} \left[\frac{1}{\sigma} \lambda \left(\frac{T_{ij} - \gamma^T Z^{(i)}}{\sigma} \right) \right]^{\delta_{ij}} R \left(\frac{T_{ij} - \gamma^T Z^{(i)}}{\sigma} \right) \right] \pi(\gamma, \sigma)}{K_I} \quad (4.4)$$

avec $K_I =$

$\int_0^\infty \left\{ \int_0^\infty \cdots \int_0^\infty \left[\prod_{i=1}^k \prod_{j=1}^{n_i} \left[\frac{1}{\sigma} \lambda \left(\frac{T_{ij} - \gamma^T Z^{(i)}}{\sigma} \right) \right]^{\delta_{ij}} R \left(\frac{T_{ij} - \gamma^T Z^{(i)}}{\sigma} \right) \right] \pi(\gamma, \sigma) d\gamma_1 \cdots d\gamma_m \right\} d\sigma$, une constante de normalisation qui assure la propriété d'une densité de probabilité pour la fonction $\pi_{apo}(\gamma, \sigma | T)$.

4.2.1.2 Estimation bayésienne des paramètres de modèles

Nous proposons deux approches pour estimer les paramètres inconnus des modèles SVA :

1. la méthode classique consistant à définir les estimateurs ponctuels à partir des premiers moments des fonctions marginales associées aux différents paramètres et les intervalles de confiance par les estimations des variances obtenues par les deuxièmes moments,
2. la méthode MAP (Maximum of A Priori) qui détermine les estimateurs ponctuels à partir des modes des fonctions marginales et les intervalles de confiance par les estimations des variances par la matrice de Fisher.

méthode classique En faisant l'hypothèse que les variables γ_i et σ sont indépendantes, alors il est possible de définir les fonctions de densité marginale a posteriori $g(\gamma_i)$ et $g(\sigma)$ des paramètres γ_i et σ par :

$$g_i(\gamma_i) = \int_{-\infty}^0 \int \cdots \int \pi_{apo}(\gamma, \sigma | T) d\gamma_j d\sigma \quad (4.5)$$

et

$$g(\sigma) = \int \cdots \int \pi_{apo}(\gamma, \sigma | T) d\gamma_i \quad (4.6)$$

Dans ce contexte, les estimateurs ponctuels sont les moments initiaux d'ordre 1 :

$$\hat{\gamma}_i = \int_{\Omega_{\gamma_i}} \gamma_i g_i(\gamma_i) d\gamma_i \quad (4.7)$$

et

$$\hat{\sigma} = \int_{\Omega_{\sigma}} \gamma_i g(\sigma) d\sigma \quad (4.8)$$

Les variances sont définies par les moments centrés d'ordre 2 :

$$V(\hat{\gamma}_i) = \int_{\Omega_{\gamma_i}} g_i(\gamma_i) (\gamma_i - \hat{\gamma}_i)^2 d\gamma_i \quad (4.9)$$

et

$$V(\hat{\sigma}) = \int_{\Omega_{\sigma}} g(\sigma) (\sigma - \hat{\sigma})^2 d\sigma \quad (4.10)$$

La méthode nécessite de recourir à des méthodes numériques. Les variances sont utilisées pour faire les estimations des intervalles de confiance selon les relations 2.37 et 2.38. Il est possible également de définir l'intervalle de confiance bilatéral symétrique au niveau de confiance $1 - \alpha$ tel que :

$$\text{prob}(\gamma_i^{\min} \leq \gamma_i \leq \gamma_i^{\max}) = 1 - \alpha \quad (4.11)$$

avec γ_i^{\min} et γ_i^{\max} solutions de :

$$\int_{\Omega_{\gamma_i}}^{\gamma_i^{\min}} g_i(\gamma_i) d\gamma_i = \alpha/2 \quad (4.12)$$

$$\int_{\Omega_{\gamma_i}}^{\gamma_i^{\max}} g_i(\gamma_i) d\gamma_i = 1 - \alpha/2 \quad (4.13)$$

On procède de la même manière pour le paramètre σ .

méthode MAP (Maximum of A Priori) Cette méthode considère la fonction de densité a posteriori $\pi_{apo}(\gamma, \sigma | T)$ et les estimateurs ponctuels des paramètres inconnus (γ, σ) sont estimés de telle sorte qu'ils maximisent :

$$(\hat{\gamma}, \hat{\sigma}) = \text{Arg max}(\pi_{apo}(\gamma, \sigma | T)). \quad (4.14)$$

Comme dans le cas de la méthode MV classique, ces estimateurs sont obtenus par le calcul des dérivées à l'ordre 1 de la logvraisemblance qui sont :

$$\begin{aligned}
U_l(\gamma, \sigma) &= \frac{\partial (\ln [\pi_{apo}(\gamma, \sigma | T)])}{\partial \gamma_l} = \frac{\partial \ln \left[\frac{L(\gamma, \sigma) \pi(\gamma, \sigma)}{K_I} \right]}{\partial \gamma_l} \\
&= \frac{\partial \ln [L(T | \gamma, \sigma)]}{\partial \gamma_l} + \frac{\partial \ln [\pi(\gamma, \sigma)]}{\partial \gamma_l} - \frac{\partial \ln [K_I]}{\partial \gamma_l} \\
&= \frac{\partial \ln [L(T | \gamma, \sigma)]}{\partial \gamma_l} + \frac{d \ln [\pi(\gamma, \sigma)]}{d \gamma_l} \\
&= \frac{1}{\sigma} \sum_{i=1}^k z_{il} \sum_{j=1}^{n_i} a_{ij}(\gamma, \sigma) + \frac{b-1}{\gamma_l} + \frac{d \ln [\pi(\gamma, \sigma)]}{d \gamma_l}, \quad (l = 1, \dots, m) \\
U_{m+1}(\gamma, \sigma) &= \frac{\partial (\ln [\pi_{apo}(\gamma, \sigma | T)])}{\partial \sigma} = \frac{\partial \ln \left[\frac{L(\gamma, \sigma) \pi(\gamma, \sigma)}{K_I} \right]}{\partial \sigma} \\
&= \frac{\partial \ln [L(T | \gamma, \sigma)]}{\partial \sigma} + \frac{\partial \ln [\pi(\gamma, \sigma)]}{\partial \sigma} - \frac{\partial \ln [K_I]}{\partial \sigma} \\
&= \frac{\partial \ln [L(T | \gamma, \sigma)]}{\partial \sigma} + \frac{d \ln [\pi(\gamma, \sigma)]}{d \sigma} \\
&= \frac{1}{\sigma} \sum_{i=1}^k \sum_{j=1}^{n_i} \{v_{ij}(\gamma, \sigma) a_{ij}(\gamma, \sigma) - \delta_{ij}\} + \frac{d-1}{\sigma} + \frac{d \ln [\pi(\gamma, \sigma)]}{d \sigma}
\end{aligned} \tag{4.15}$$

avec :

$$\begin{aligned}
v_{ij}(\gamma, \sigma) &= \frac{T_{ij} - \gamma^T Z^{(i)}}{\sigma}, \\
a_{ij}(\gamma, \sigma) &= \lambda(v_{ij}(\gamma, \sigma)) - \delta_{ij} (\ln \lambda)'(v_{ij}(\gamma, \sigma)), \\
\frac{\partial \ln [\pi_\sigma(\gamma, \sigma)]}{\partial \sigma} &= \frac{d \left(\ln \left[d \left(\frac{\sigma}{c} \right)^d + \ln [\pi(\gamma, \sigma)] \right] \right)}{d \sigma}, \\
\frac{\partial \ln [\pi_{\gamma_l}(\gamma, \gamma_l)]}{\partial \gamma_l} &= \frac{d \left(\ln \left[b \left(\frac{\gamma_l}{a} \right)^b + \ln [\pi(\gamma, \sigma)] \right] \right)}{d \gamma_l}.
\end{aligned}$$

Les estimateurs de maximum d'a posteriori $\hat{\sigma}, \hat{\gamma}$ peuvent être obtenus en résolvant le système d'équations :

$$U_l(\gamma, \sigma) = 0 \quad (l = 1, \dots, m+1) \tag{4.16}$$

et l'estimateur de la fonction de fiabilité R_{X_0} est :

$$\hat{R}_{X_0}(t) = R \left(\frac{\ln t - \hat{\gamma}^T Z^{(0)}}{\hat{\sigma}} \right) \tag{4.17}$$

L'information de Fisher, dans ce cas, respecte le principe de l'équation 2.26 et s'applique à la fonction qui décrit l'information apportée sur les paramètres, $\pi_{apo}(\gamma, \sigma | T)$, [Manry et al., 1996] :

$$\begin{aligned}
I^{MAP}(\gamma, \sigma) &= E \left[\left(\frac{\partial \log \pi_{apo}(\gamma, \sigma | T)}{\partial \gamma \partial \sigma} \right)^2 \middle| \hat{\gamma}, \hat{\sigma} \right] \\
&= E \left[\left(\frac{\partial \log L(T | \gamma, \sigma)}{\partial \gamma \partial \sigma} \right)^2 \middle| \hat{\gamma}, \hat{\sigma} \right] + E \left[\left(\frac{\partial \log \pi(\gamma, \sigma)}{\partial \gamma \partial \sigma} \right)^2 \middle| \hat{\gamma}, \hat{\sigma} \right] \\
&= I^L(\gamma, \sigma) + I_{\pi}^{AP}(\gamma, \sigma)
\end{aligned} \tag{4.18}$$

avec :

$$\begin{aligned}
I^L(\gamma, \sigma) &\quad \text{définie dans les équations 3.29 – 3.31 et} \\
I_{\pi}^{AP}(\gamma, \sigma) &= E \left[\left(\frac{\partial \log \pi(\gamma, \sigma)}{\partial \gamma \partial \sigma} \right)^2 \middle| \hat{\gamma}, \hat{\sigma} \right].
\end{aligned}$$

On observe rapidement que l'information apportée sur les paramètres γ et σ augmente par rapport à l'approche classique avec $I_{\pi}^{AP}(\gamma, \sigma)$. Ceci implique la diminution des variances/covariances des estimateurs ponctuels des paramètres γ et σ , (donnés par l'inégalité Cramer-Rao (équation 2.30)).

En reprenant les équations 3.29 - 3.34, pour calculer l'intervalle de confiance $(1-\alpha)$ sur l'estimateur de la fonction de fiabilité, on obtient la matrice de Fisher $I^{MAP}(\gamma, \sigma)$ avec les termes suivants:

$$\begin{aligned}
I_{ls}^{MAP}(\gamma, \sigma) &= I_{ls}^L(\gamma, \sigma) + \left[-\frac{\partial^2 \ln \pi(\gamma, \sigma)}{\partial \gamma_l \partial \gamma_s} \right] \\
I_{l,m+1}^{MAP}(\gamma, \sigma) &= I_{l,m+1}^L(\gamma, \sigma) + \left[-\frac{\partial^2 \ln \pi(\gamma, \sigma)}{\partial \gamma_l \partial \sigma} \right] \\
I_{m+1,m+1}^{MAP}(\gamma, \sigma) &= I_{m+1,m+1}^L(\gamma, \sigma) + \left[-\frac{\partial^2 \ln \pi(\gamma, \sigma)}{\partial \sigma^2} \right]
\end{aligned} \tag{4.19}$$

avec :

$$\begin{aligned}
\frac{\partial^2 \ln \pi(\gamma, \sigma)}{\partial \gamma_l \partial \gamma_s} &= \begin{cases} 0 & \text{pour } l \neq s \\ -\frac{2(-1+b)}{\gamma_l^2} & \text{pour } l = s, \end{cases} \\
\frac{\partial^2 \ln \pi(\gamma, \sigma)}{\partial \sigma \partial \gamma_l} &= 0, \\
\frac{\partial^2 \ln \pi(\gamma, \sigma)}{\partial \sigma^2} &= -\frac{2(-1+d)}{\sigma^2}
\end{aligned}$$

Nous avons considéré une indépendance totale entre les variables (équation ??).

De la même manière que pour le cas non-Bayésien, on observe que l'estimateur \hat{R} est une fonction régulière de γ et σ . On va définir :

$$Q_{X_0}(t) = \ln \frac{R_{X_0}(t)}{1 - R_{X_0}(t)} \quad \text{et} \quad \hat{Q}_{X_0}(t) = \ln \frac{\hat{R}_{X_0}(t)}{1 - \hat{R}_{X_0}(t)} \tag{4.20}$$

La fonction $Q_{X_0}(t)$ prend ses valeurs dans \mathfrak{R} donc la vitesse de convergence de \hat{Q}_{X_0} vers la loi limite est plus grande que celle de \hat{R}_{X_0} vers sa loi limite, et on obtient que la loi de

$$\frac{\hat{Q}_{X_0}(t) - Q_{X_0}(t)}{\hat{\sigma}_{Q_0}}$$

est approximée par la loi normale centrée réduite $N(0, 1)$, avec :

$$\begin{aligned} (\hat{\sigma}_{Q_0})^2 &= \left(\frac{\partial \hat{Q}_{X_0}}{\partial \hat{\gamma}_0}, \dots, \frac{\partial \hat{Q}_{X_0}}{\partial \hat{\gamma}_m}, \frac{\partial \hat{Q}_{X_0}}{\partial \hat{\sigma}} \right) [I^{MAP}(\gamma, \sigma)]^{-1} \left(\frac{\partial \hat{Q}_{X_0}}{\partial \hat{\gamma}_0}, \dots, \frac{\partial \hat{Q}_{X_0}}{\partial \hat{\gamma}_m}, \frac{\partial \hat{Q}_{X_0}}{\partial \hat{\sigma}} \right) \\ &= \left(\frac{R'(R^{-1}(\hat{R}_{X_0}(t)))}{\hat{\sigma}^2 \hat{R}_{X_0}(t) (1 - \hat{R}_{X_0}(t))} \right)^2 \times \left[\hat{\sigma}^4 \sum_{l=0}^m \sum_{s=0}^m z_{0l} z_{0s} I_{MAP}^{ls}(\hat{\gamma}, \hat{\sigma}) \right. \\ &\quad \left. - 2\hat{\sigma}^2 \hat{\gamma}^T Z^{(0)} \sum_{l=0}^m z_{0l} I_{MAP}^{l,m+1}(\hat{\gamma}, \hat{\sigma}) + (\hat{\gamma}^T Z^{(0)})^2 I_{MAP}^{m+1,m+1}(\hat{\gamma}, \hat{\sigma}) \right] \end{aligned} \quad (4.21)$$

avec $I_{MAP}^{l,s}$ les termes de la matrice $[I^{MAP}(\gamma, \sigma)]^{-1}$.

L'intervalle de confiance approximatif $(1-\alpha)$ pour $\hat{R}_{X_0}(t)$ est:

$$\left(1 + \frac{1 - \hat{R}_{X_0}(t)}{\hat{R}_{X_0}(t)} \exp \left\{ \pm \hat{\sigma}_{Q_0} w_{1-\alpha/2} \right\} \right)^{-1} \quad (4.22)$$

avec w_α le α -quantile de la loi $N(0, 1)$.

4.3 L'information a priori

L'introduction d'une loi a priori $\pi(\theta)$ sur les paramètres divise toujours les statisticiens. Le passage de la notion d'un paramètre inconnu à celle d'un paramètre aléatoire représente un saut que beaucoup se refusent à effectuer [Robert, 1992]. Si, dans certains modèles, le paramètre à estimer résulte de l'action de multiples facteurs en amont et peut donc effectivement être perçu comme partiellement aléatoire (on pourrait ainsi citer la physique quantique et sa représentation probabiliste du monde), il est d'autres modèles où la perception du paramètre comme produit d'une expérience aléatoire semble impossible. Ainsi, les adversaires de l'approche Bayésienne avancent l'exemple de l'estimation de quantités physiques, telle la vitesse de la lumière, c . On peut rétorquer à cette objection que la véritable valeur de c n'est jamais connue, du fait de la limitation des appareils de mesure, et qu'il est donc légitime de considérer c comme une variable uniforme dans $[c_0 - \epsilon, c_0 + \epsilon]$, ϵ étant la précision maximale des instruments de mesure et c_0 la mesure retenue. Nous soulignons que l'importance de la loi a priori n'est pas tant dans le fait que le paramètre puisse ou ne puisse pas être conceptuellement envisagé comme la réalisation d'une variable aléatoire, que dans le fait que $\pi(\bullet)$ représente un moyen efficace (le seul?) de résumer l'information a priori disponible sur les paramètres, ainsi que l'incertitude sur la valeur de cette information, et donc de permettre une appréciation quantitative de ces paramètres. C'est également l'unique manière cohérente de raisonner conditionnellement (au sens probabiliste) aux observations [Robert, 2001].

L'utilisation d'une loi $\pi(\theta)$ implique trois niveaux de discussion :

- La source : quelle est la source d'information
- La forme : quelle est le choix de la loi
- Les paramètres : quelles sont les valeurs des paramètres à choisir

4.3.1 Sources d'information

Il existe plusieurs sources d'informations où sont synthétisées toute la connaissance disponible sur les paramètres influençant la fiabilité des composants. Nous distinguerons, dans le cadre de cette thèse, deux sources distinctes :

- recueils de données issus d'analyse de REX,
- avis d'expert.

4.3.1.1 Recueils de données

Le recueil des données est un ensemble de données validées et/ou élaborées après un long processus d'expertise et de traitement, relatif à un domaine de connaissance et organisé pour être offert aux consultations d'utilisateurs.

En électronique, un domaine dans lequel le calcul de la fiabilité est pratiqué depuis de nombreuses années, les bases de données de fiabilité sont disponibles et nombreuses. En mécanique l'utilisation des recueils de données est plus récente. Il y a une certaine difficulté à constituer ces recueils, compte tenu de la complexité des composants. Néanmoins, ils sont de plus en plus utilisés.

En pratique, on utilise souvent des bases de données connues ou, mieux encore, quand cela est possible, les données issues de retours d'expériences auprès des fabricants des composants. Pour les composants mécaniques ces données sont difficiles à obtenir parce qu'elles ne sont pas communiquées ou publiées et, dans la plupart des cas, elles sont exploitées confidentiellement.

Ces recueils sont mis à jour régulièrement pour tenir compte des évolutions technologiques. En électronique, la durée de vie des données fournies par les recueils est relativement courte (de 3 à 6 ans) ; elle est plus longue en mécanique (de 30 à 40 ans). En électronique, les recueils des données sont souvent exhaustifs sachant que les composants suivent une loi exponentielle. En mécanique, l'hypothèse du taux de défaillance constant est rarement justifiée.

Dans la majorité des recueils, les informations disponibles sont la dénomination du composant, la moyenne des temps de bon fonctionnement (MTTF), le taux de défaillance moyen ou calculé dans l'hypothèse selon laquelle il est constant, l'intervalle de confiance associé et un coefficient de pondération du taux de défaillance dépendant de l'environnement. Le tableau 4.1 est un extrait d'un tableau du recueil NPRD-3 [NPRD-3, 1985] avec les données suivantes sur les composants : le taux de défaillance (caractérisé par les valeurs minimale, moyenne et maximale), le nombre des défaillances et les temps cumulés d'exécution. Ainsi dans le même tableau, nous observons que le nombre de défaillances est directement lié à l'intervalle de confiance, un faible nombre de défaillances implique un grand intervalle de confiance et vice-versa [Mihalache, 2007].

| Composant | λ_{moy} $10^{-6}h^{-1}$ | Limite inf. 20% | Limite sup 20% 20% | Nbr. de défaillances | Temps cumulé h^{-6} |
|------------|------------------------------------|--------------------|-----------------------|-------------------------|--------------------------|
| soupape | 214.286 | 109.372 | 396.948 | 3 | 0.014 |
| Capteur | 672.367 | 651.597 | 693.921 | 747 | 1.111 |
| Actionneur | 1111.11 | 1053.916 | 1171.927 | 270 | 0.243 |
| Led | 0.267 | 0.264 | 0.289 | 363 | 1312.883 |

Table 4.1 – Extrait du recueil NPRD-3

Un autre exemple traite le lien entre les modes de défaillances et les valeurs d'énergie d'activation. La figure 4.3 présente un extrait d'un recueil dans le domaine de l'électronique.

Dans le domaine du logiciel il y a peu de recueils de données. Les erreurs sont introduites lors de la conception ou bien lors de la programmation. Microsoft a estimé qu'un programmeur introduit environ 6 défaillances sur 1000 lignes de code [Lyonnet, 2006], [MIL-HDBK-338B, 1998].

Les données collectées permettent d'établir des modèles de fiabilité prévisionnelle pour les composants mécaniques, électroniques et logiciel.

4.3.1.2 L'avis d'experts

En sûreté de fonctionnement comme dans d'autres domaines, le jugement d'expert peut être considéré comme l'une des réponses possibles à un problème technique, un appoint à un retour d'expérience déficient ou un support d'aide à la décision. Il est en effet utilisé pour compléter ou suppléer les données objectives éventuellement disponibles sur le problème considéré lorsque ces données sont peu nombreuses, dispersées voire inexistantes ou lorsqu'elles sont douteuses ou inapplicables. C'est bien souvent la seule source d'information disponible pour effectuer un état de l'art sur un sujet difficile, complexe, nouveau ou peu compris [Poncet, 2005].

Dans le domaine technique, l'expertise peut être considérée comme une source de données, souvent qualitative, quelquefois quantitative. Mais comme toute donnée, les données d'expertise sont une représentation imparfaite et incertaine de ce qu'elles sont censées représenter. La qualité de ces données dépend essentiellement :

- de l'information disponible et de l'expert et, en particulier, de sa formation : son éducation, sa culture, sa profession, sa formation théorique et pratique sur le sujet, son expérience, son

| Mécanismes de défaillance | Causes de défaillance | Mode de défaillance | Etapes de fabrication | Types de défaillance | E_a , eV | Modes accélér. | Types de défaut |
|--|---|--------------------------------|--------------------------------------|--------------------------------|-------------|----------------------------|-----------------|
| 1) Contamination ionique | Présence d'ions alcalins dans oxyde de grille | Dérive de la tension de seuil | Photoresist. Passivation | Jeunesse Vieillesse | 1 1,4 | E, T | D |
| 2) Charges de surface | Présence de charges en surfaces des oxydes de grille | Dérive de la tension de seuil | Diffusion Trait. à chaud Passivation | Vieillesse | 0,5 1 | E, T | D |
| 3) Inversion de charges | Charges induites | Dérive de la tension de seuil | Passivation | Vieillesse | 1 | E, T | D |
| 4) Accumulation de charges en surface | Impuretés ioniques dans la résine | Dérive de la tension de seuil | Encapsulation | | 1 1,35 | E, T | D |
| 5) Perte de charges | Perte de charges entraînant une dérive de la tension de seuil | Mauvais fonctionnement mémoire | | Jeunesse Aléatoires Vieillesse | 0,6 1,4 | T | D |
| 6) Polarisation | Présence de molécules polarisables dans oxyde de grille | Dérive de la tension de seuil | Dopage phosphore Passivation | Vieillesse | 1 | E, T | D |
| 7) Défauts d'oxyde | Cassures, trous fêlures dans oxyde | Circuits ouverts | Croissance oxyde | Jeunesse Aléatoires | 0,3 0,5 | E, T | C |
| 8) Porteurs chauds | Ions piégés dans l'oxyde | Dérive de la tension de seuil | Inadéquation conception procédés | Vieillesse | -1 -1,4 | $T < 0^\circ\text{C}$ V | D |
| 9) Claquage diélectrique | ESD, surtension | Circuits ouverts Fuites | Photoresist Etching | Jeunesse Aléatoires | 0,2 1 | E, T | C |
| 10) Interaction (Al-SiO ₂) | Destruction de l'oxyde | Court-circuits | Croissance oxyde | Vieillesse | 1,77 2,5 | E, T | C |

Figure 4.3 – Des valeurs pour l'énergie d'activation par rapport aux modes de défaillances

environnement, sa personnalité, et d'une façon générale de la qualité intrinsèque de l'expert, mais aussi de sa motivation et de son sérieux, de sa façon de recueillir, appréhender, transmettre, corriger et agréger ses connaissances,

- de la complexité relative du problème et de l'état général des connaissances sur ce problème, et de la manière dont cette complexité est modélisée dans l'esprit de l'expert,
- de la façon dont va s'instaurer le dialogue entre l'expert et l'analyste chargé de recueillir puis d'interpréter son expertise.

D'après Meyer et Booker [MEY93], est considéré comme un expert dans un domaine, toute personne ayant une connaissance particulière avec le niveau de détail nécessaire à l'étude de ce domaine.

L'inférence Bayésienne peut se porter sur l'avis d'un ou des plusieurs experts.

4.3.2 Le choix de la loi a priori

Le choix d'une distribution a priori est une démarche complexe et peut laisser place à une appréciation subjective. Ce choix dépend principalement de la connaissance préalable, de la qualité des informations disponibles sur le paramètre étudié.

La distribution a priori doit intégrer, d'une part, toutes les connaissances disponibles sur le paramètre étudié et, d'autre part, toutes les incertitudes liées à ce paramètre. En fonction de la connaissance disponible, nous pouvons définir les principales distributions susceptibles

de caractériser la distribution a priori. Si la connaissance initiale est faible, l'a priori sera une distribution simple, peu informative. Si tout paramètre cherché a la même probabilité d'être choisi, l'a priori suit une distribution uniforme. Si la connaissance initiale est solide, l'a priori sera une distribution plus sophistiquée.

Concernant la modélisation effective de la loi $\pi(\theta)$ a priori à partir de la connaissance a priori, de nombreuses méthodologies ont été développées pour modéliser la connaissance a priori à partir du type d'informations disponibles.

Plusieurs principes peuvent être appliqués :

- le maximum d'entropie : la loi a priori $\pi(\theta)$ maximisant l'entropie correspond en effet à la loi minimisant l'information a priori apporté sur θ ,
- les méthodes Bayésiennes empiriques : utilisent les observations et la loi marginale pour estimer les paramètres de la loi a priori,
- les modèles hiérarchiques : modélisent le manque d'information sur les paramètres de la loi a priori par une nouvelle distribution sur ces paramètres, etc.

Même dans les situations où l'information a priori est disponible, il est rare de pouvoir proposer une détermination exacte de la loi a priori. Il importe ainsi de pouvoir mesurer les conséquences de cette imprécision sur l'inférence a posteriori, afin de contrôler que l'apport arbitraire dans la loi a priori n'est pas prédominant. Cette étude est dite analyse de sensibilité ou de robustesse. Nous proposons d'utiliser les pseudo-lois conjuguées. En effet, pour étendre le concept d'indifférence au choix de la meilleure connaissance a priori en fonction des données observées, Clarotti démontre dans [Clarotti, 1998b] et [Procaccia, 1998] l'intérêt particulier des lois conjuguées.

4.3.2.1 Lois conjuguées

Pour des raisons de calcul purement théorique, la distribution a priori dépend de la famille mathématique de la fonction de vraisemblance; cette dernière condition n'est pas obligatoire pour une résolution numérique.

En fait, le problème consiste à définir quelle densité a priori sera susceptible de laisser le pouvoir le plus cohérent entre les données et la connaissance a priori. Il est donc nécessaire de trouver une série de densités de probabilités a priori qui permette de déterminer avec bon sens quel pouvoir de décision sera laissé aux données.

Il est ainsi démontré que les familles de lois dites conjuguées ne se contentent pas uniquement de simplifier les calculs mais surtout que ce type de lois maximise le pouvoir des données observées.

Le tableau 4.2 donne les principales lois conjuguées utilisées - [Procaccia et al., 1992b], [Suhner, 1994]

La recherche de ces lois peut être problématique car nous ne disposons pas des expressions analytiques pour définir la loi a priori conjuguée d'une vraisemblance définie par un modèle SVA. Néanmoins, nous proposons de trouver une loi a priori par simulation vérifiant les propriétés de conjugaison.

4.3.2.2 Détermination de la loi a priori pseudo conjuguée par simulation

Pour obtenir la forme de la distribution a priori qui soit la conjuguée bayésienne d'une vraisemblance donnée, il est nécessaire de procéder en deux étapes :

| $f(x \theta)$ | $\pi(\theta)$ | $\pi_{apo}(\theta x)$ |
|--|---|---|
| Normale $N(\theta, \sigma^2)$ | Normale $N(\mu, \tau^2)$ | $N(\rho(\sigma^2\mu + r^2x), \rho\sigma^2r^2)$ avec $\rho^{-1} = \sigma^2 + r^2$ |
| Poisson $P(\theta)$ | Gamma $G(\alpha, \beta)$ | $G(\alpha + x, \beta + 1)$ |
| Gamma $G(v, \theta)$ | Gamma $G(\alpha, \beta)$ | $G(\alpha + v, \beta + x)$ |
| Binomiale $B(n, \theta)$ | Bêta $Be(\alpha, \beta)$ | $Be(\alpha + x, \beta + n - x)$ |
| Négative Binomiale $Neg(m, \theta)$ | Bêta $Be(\alpha, \beta)$ | $Be(\alpha + m, \beta + x)$ |
| Multinomiale $M_k(\theta_1, \dots, \theta_k)$ | Dirichlet $D(\alpha_1, \dots, \alpha_k)$ | $D(\alpha_1 + x_1, \dots, \alpha_k + x_k)$ |
| Bêta $Be(\alpha, \theta)$ | Exponentielle $Exp(\lambda)$ | $exp(\lambda - \log(1 - x))$ |
| Normale $N(\mu, 1/\theta)$ | Gamma $G(\alpha, \beta)$ | $G(\alpha + 0.5, \beta + (\mu - x)^2/2)$ |

Table 4.2 – Lois conjuguées naturelles des familles exponentielles classiques

- la première consiste à déterminer la distribution a posteriori en considérant un a priori non informatif :

$$\Pi_{apos1}(\theta|t_i) = \frac{L(t_1, \dots, t_n|\theta) \cdot \Pi_0(\theta)}{\int_{\Omega} L(t_1, \dots, t_n|\theta) \cdot \Pi_0(\theta) \cdot d\theta}$$

- dans un deuxième temps, nous construisons la distribution a priori $\Pi_1(\theta)$ selon la forme définie par $\Pi_{apos1}(\theta|t_i)$ et ensuite nous déterminons la distribution a posteriori :

$$\Pi_{apos2}(\theta|t_i) = \frac{L(t_1, \dots, t_n|\theta) \cdot \Pi_1(\theta)}{\int_{\Omega} L(t_1, \dots, t_n|\theta) \cdot \Pi_1(\theta) \cdot d\theta}$$

Nous vérifions que le type de distribution de $\Pi_{apos2}(\theta|t_i)$ est identique à $\Pi_1(\theta)$. Si cela est le cas, ce type respecte les conditions de conjugaison bayésienne.

Dans le cas des modèles SVA, il est difficile de traiter ces deux étapes analytiquement. Aussi, nous proposons de déterminer approximativement les lois conjuguées bayésiennes par simulation.

La procédure des simulations est la suivante :

- on fixe le plan d'essai (niveau d'accélération, taille des échantillons à tester sous chaque niveau de stress) permettant de construire la vraisemblance,
- à l'aide des informations a priori sur les paramètres, on cherche les valeurs les plus probables associées aux paramètres [Poncet et al., 2004],
- on construit la simulation de Monte-Carlo en répétant (n fois) les étapes suivantes :
 - on simule des instants de défaillance aléatoires sur les hypothèses estimées au point 2,
 - on calcule les estimations des paramètres par MV.

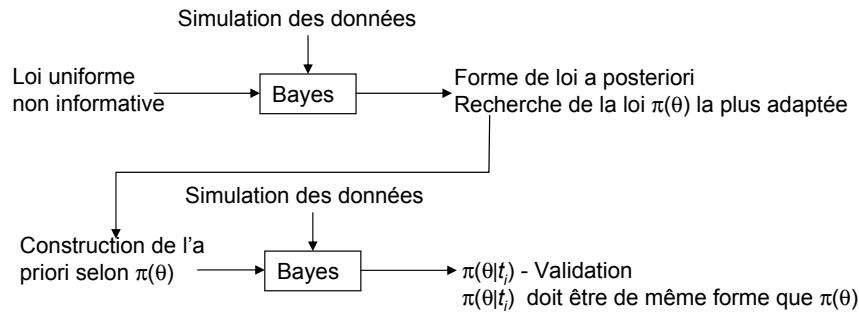


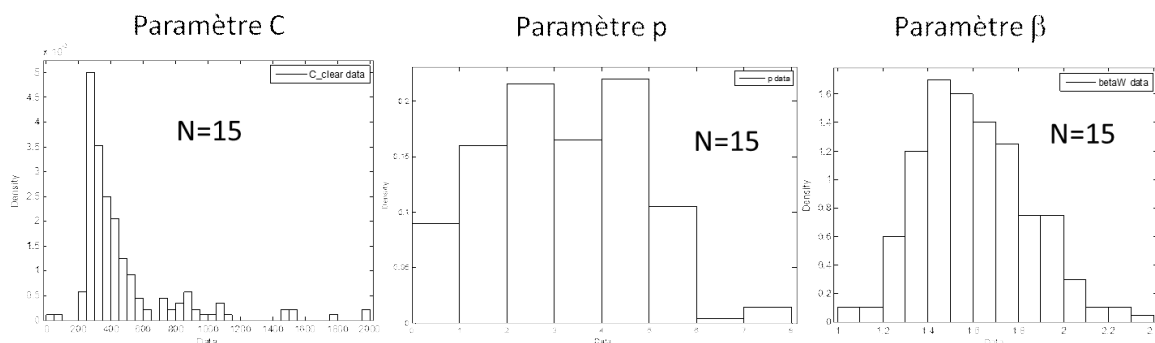
Figure 4.4 – Recherche d’une pseudo loi conjuguée

4. on construit les histogrammes des estimateurs MV,
5. on fait des hypothèses sur les formes des lois les plus proches des histogrammes et on vérifie la validité de ces hypothèses (test du Chi-Deux, test de Kolmogorov Smirnov, ...).
6. on construit les lois a priori selon les formes obtenues dans l’étape 5 à partir des informations de l’étape 2,
7. on réitère les étapes 3, 4 et 5.
8. on vérifie que les formes des distributions obtenues dans l’étape 7 sont identiques à celles de l’étape 5.

La construction d’a priori serait faite en respectant la forme de lois validée au point 8.

La figure 4.4 synthétise la procédure.

Une fois la modélisation a priori réalisée et les décisions prises, concernant le type de loi, il reste un point à traiter. Il faut en effet vérifier que la loi a priori déterminée est la plus pertinente et, notamment, qu’elle n’influence pas la vraisemblance de manière à dégrader le résultat. Cette partie constitue un élément peu étudié à ce jour. Les figures 4.5 et 4.6 montrent que la forme des distributions a posteriori des paramètres d’un modèle SVA paramétrique (Weibull-Puissance Inverse défini dans l’exemple sur le roulement à billes de la section 3.3) de paramètres C , p et β ne changent pas de forme notablement avec la taille d’échantillon ($n = 15$ et $n = 50$).

Figure 4.5 – Application pour un Modèle SVA Weibull-Puissance inverse avec $n=15$

Il reste à déterminer les paramètres des lois a priori à partir des informations disponibles.

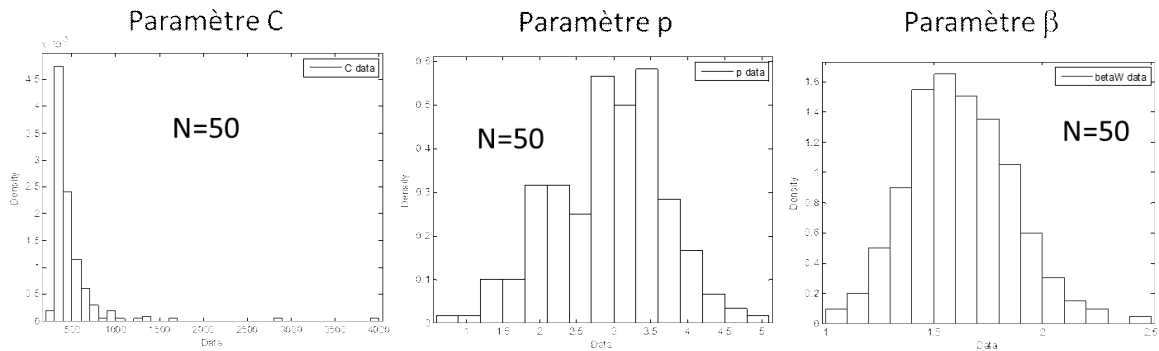


Figure 4.6 – Application pour un Modèle SVA Weibull-Puissance inverse avec n=50

4.3.3 Construction de la loi a priori

Une fois que l’on a choisi les types de loi à utiliser pour modéliser la distribution a priori, il reste ensuite à déterminer les paramètres de lois à partir des différentes informations disponibles. Rappelons que l’n fait l’hypothèse que tous les paramètres de modèles sont indépendants qui implique de la distribution a priori est le produit des distributions a priori individuelles associées aux différents paramètres ($\Pi(\theta) = \prod_{i=1}^q \Pi_i(\theta_i)$). En fonction du nombre de paramètres de la loi à modéliser, il faudra disposer d’un ou de plusieurs types d’informations.

En fonction des informations fournies, plusieurs méthodes sont disponibles pour estimer les paramètres de lois [Suhner, 1994], [Finetti, 1974] et [Goldstein, 1999] :

- la méthode séquentielle,
- la méthode des échantillons équilibrés,
- la méthode des moments pondérés,
- la méthode des doubles moments,
- la méthode des similitudes comme l’ADEPT ou autre
- la méthode du maximum d’entropie,
- les techniques de transformation de groupe,
- la méthode des moments (moyenne et écart - type),
- la méthode des fractiles.

La plupart des sources d’informations donne celles-ci sous forme d’intervalle. Aussi, nous proposons dans la suite de la thèse d’utiliser la méthode des moments. Le principe consiste à déterminer la moyenne et la variance de loi uniforme informative bornée par les limites d’intervalle. Ainsi, lorsque l’on dispose d’une information a priori, sur le paramètre θ_i , par un intervalle $\theta_i \in [\theta_i^{min}, \theta_i^{max}]$, il est possible d’estimer la moyenne et la variance par :

$$\bar{\theta}_i = \frac{\theta_i^{min} + \theta_i^{max}}{2} \quad \text{et} \quad V(\theta_i) = \frac{(\theta_i^{max} - \theta_i^{min})^2}{12} \tag{4.23}$$

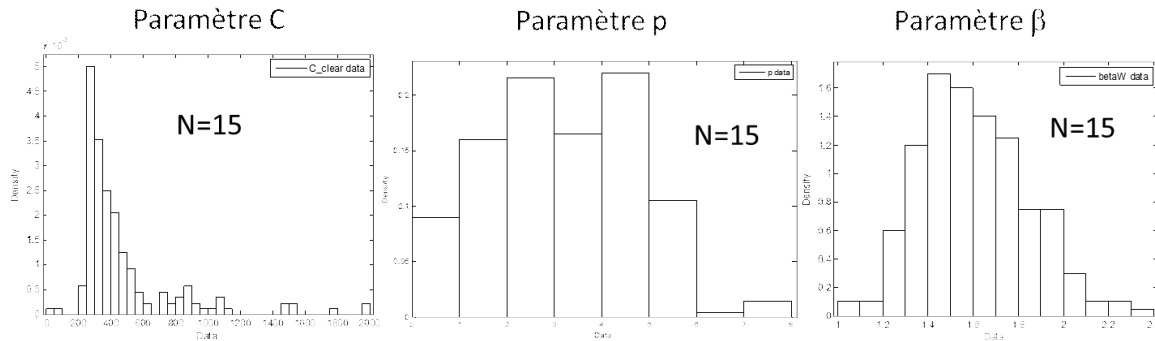
Ensuite, nous choisissons un type de loi a priori et estimons ses paramètres permettant de respecter la moyenne et la variance obtenues ci-dessus.

Pour illustrer cette méthode, considérons le cas d’un modèle SVA paramétrique (modèle Weibull-Puissance Inverse présenté précédemment pour l’exemple des roulements à billes) pour lequel nous disposons des informations a priori (tableau 4.3).

| paramètre | limite inférieure | limite supérieure |
|-----------|-------------------|-------------------|
| C | 300 | 400 |
| p | 2 | 4 |
| β | 2 | 3 |

Table 4.3 – Connaissances a priori

Les figures 4.7 et 4.8 montrent le résultat de la recherche des pseudo lois conjuguées bayésiennes par simulation. Ainsi, la figure 4.7 permet d’observer que la loi Gamma est adaptée pour modéliser les distributions a posteriori sur C , p et β en considérant des distributions a priori non informatives (des tests d’hypothèse ont été appliqués pour la validation des lois).

Figure 4.7 – Distribution a posteriori avec un a priori non informatif pour une taille d’échantillon $n=15$

Ensuite, nous recherchons les paramètres α et β des lois Gamma par la méthode des moments. La moyenne et la variance pour une loi Gamma, en fonction des paramètres, sont définies par :

$$\bar{\theta}_i = \frac{\alpha}{i}\beta \quad \text{et} \quad V(\theta_i) = \frac{\alpha}{\beta^2} \quad (4.24)$$

Aussi, nous pouvons déterminer les paramètres de la loi Gamma pour θ_i par :

$$\alpha = \frac{\bar{\theta}_i^2}{V(\theta_i)} \quad \text{et} \quad \beta = \frac{\bar{\theta}_i}{V(\theta_i)} \quad (4.25)$$

Pour les paramètres du modèle SVA du tableau 4.3, nous obtenons les paramètres des lois Gamma suivants :

| paramètre | α | β |
|-----------|----------|---------|
| C | 48 | 0.12 |
| p | 27 | 9 |
| β | 75 | 30 |

Table 4.4 – Paramètres des lois Gamma a priori

Finalement, nous simulons le modèle SVA Weibull-Puissance Inverse avec la connaissance a priori définie par les trois lois Gamma précisées au dessus et une taille d’échantillon $n=15$. La

figure 4.8 montre que les distributions a posteriori sur C , p et β sont de nouveau des lois Gamma (des tests d’hypothèse ont été appliqués pour la validation des lois).

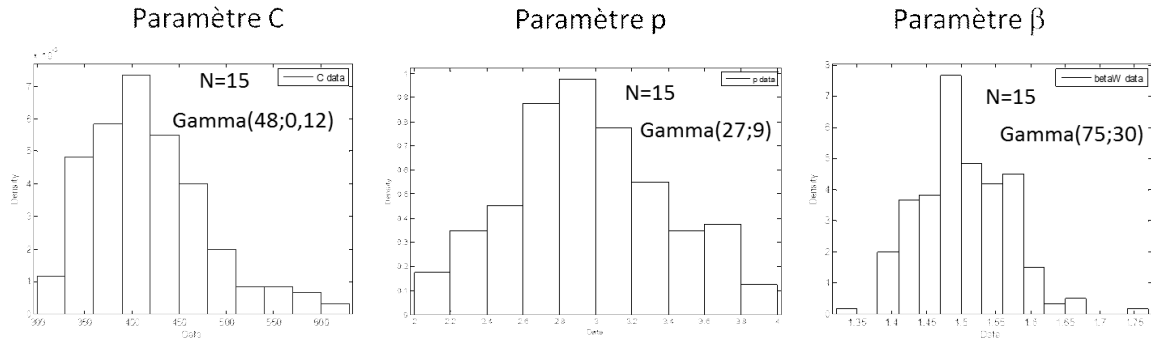


Figure 4.8 – Distributions a posteriori avec connaissance a priori

4.4 Application numérique d’estimation bayésienne d’un modèle SVA

L’exemple numérique reprend en partie les paramètres de simulation de l’exemple d’un essai accéléré sur des roulements à billes de la section 3.3. Ainsi, nous avons simulé les résultats d’un plan d’essai complet (sans censure) à deux niveaux de charge ($P_1=150$ daN et $P_2=175$ daN) et pour différentes tailles d’échantillon. Un modèle Weibull-Puissance Inverse a été utilisé pour simuler les instants de défaillance :

$$R(u, P) = e^{-\left(\frac{u}{\left(\frac{C}{P}\right)^p \left(\frac{1e6}{60 \cdot N}\right) \left(\frac{1}{0.105}\right)^{1.5}}\right)^{1.5}} \tag{4.26}$$

avec

- C est la charge dynamique de base,
- P est la charge radiale équivalente,
- $p = 3$ pour les roulements à billes et $p = 10/3$ pour les roulements à rouleaux,
- N la vitesse de rotation

Les simulations ont été réalisées avec les paramètres suivants :

- vitesse de rotation $N=1500$ tr/min,
- charge dynamique $C=400$ daN,
- paramètre de forme $\beta=1.5$,

Pour le cas Weibull-Puissance Inverse, aucune forme analytique ne permet de trouver la forme des lois conjuguées. Alors, nous avons considéré, pour des questions de simplification, que le paramètre p était fixé à 3 (valeur pour les roulements à billes) et que les distributions des paramètres C et β étaient définies par des lois normales. L’analyse se fera sur l’estimation de la probabilité de défaillance pour une durée de fonctionnement $T=100$ h et une charge nominale $P_0=125$ daN selon l’équation :

$$P_f = 1 - e^{-\left(\frac{u}{\left(\frac{C}{P}\right)^p \left(\frac{1e6}{60 \cdot N}\right) \left(\frac{1}{0.105}\right)^{1.5}}\right)^{1.5}} \quad (4.27)$$

Pour étudier l'effet aléatoire de la simulation, 200 répétitions sont réalisées pour chaque configuration de plan (taille d'échantillon $n=10, 15, 25$ et 50).

Trois cas de connaissances a priori sont présentés :

1. un cas d'un a priori cohérent (c'est-à-dire centré sur la valeur théorique) et peu informatif (variance importante),
2. un cas d'un a priori très informatif et cohérent,
3. un cas d'un a priori très informatif et incohérent (décentré par rapport à la valeur théorique).

4.4.1 A priori cohérent et peu informatif

Pour le premier cas, des distributions a priori peu informatives sur les paramètres $\theta = \{C, \beta\}$ ont été considérées.

Le tableau 4.5 donne l'opinion d'un expert sur les paramètres.

| paramètre | limite inférieure | limite supérieure |
|-----------|-------------------|-------------------|
| C | 100 | 700 |
| β | 0,5 | 2,5 |

Table 4.5 – Intervalles a priori

Nous déduisons les paramètres des loi normales a priori par la méthode des moments (tableau 4.6).

| paramètre | moyenne μ | écart-type σ |
|-----------|---------------|---------------------|
| C | 400 | 173,2 |
| β | 1,5 | 0,58 |

Table 4.6 – Paramètres des lois Normales a priori

Les résultats de simulations sont présentés dans la figure 4.9. Les courbes en trait plein et pointillé représentent, respectivement, les estimations classiques et bayésiennes.

Nous constatons par rapport aux résultats :

- l'estimateur ponctuel reste proche de la valeur théorique ce qui se justifie par le fait que les deux distributions Normales a priori sont centrées sur les valeurs théoriques,
- la taille d'échantillon, entre $n_1 = n_2 = 10$ et $n_1 = n_2 = 15$, a une influence importante sur l'estimation de l'intervalle de confiance de P_f ,
- l'estimation de l'intervalle de confiance montre une sensibilité à l'apport d'une faible information a priori.

On constate que l'intérêt de la démarche Bayésienne diminue avec l'information objective (taille d'échantillon).

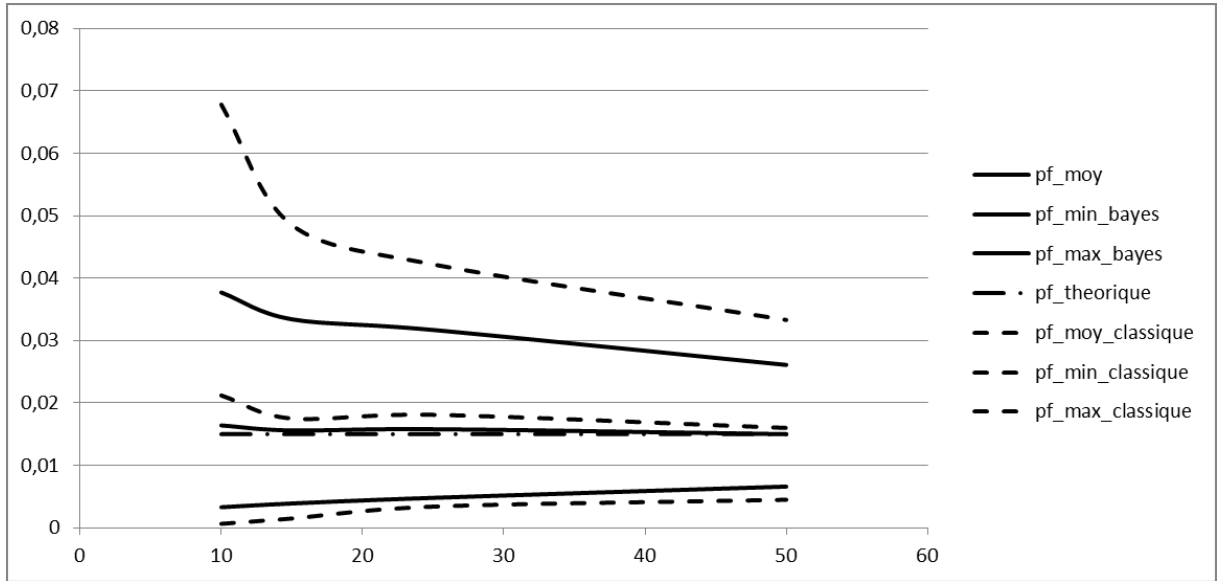


Figure 4.9 – Résultats du Monte Carlo pour les estimations classique et bayésienne de P_f avec les intervalles de confiance à 90%

4.4.2 A priori cohérent et très informatif

Pour le deuxième cas, des distributions a priori très informatives et cohérentes sur les paramètres ont été considérées.

Par une distribution a priori très informative, on comprend une distribution caractérisée par une variance très faible. Par une distribution a priori très cohérente on comprend une distribution avec une moyenne très proche de la valeur théorique.

Le tableau 4.7 donne l’opinion d’expert sur les paramètres.

| paramètre | limite inférieure | limite supérieure |
|-----------|-------------------|-------------------|
| C | 300 | 500 |
| β | 1 | 2 |

Table 4.7 – Prior intervalles

Nous déduisons les paramètres des loi normales a priori par la méthode des moments (tableau 4.8).

| paramètre | moyenne μ | écart-type σ |
|-----------|---------------|---------------------|
| C | 400 | 57,73 |
| β | 1,5 | 0,29 |

Table 4.8 – Paramètres des lois Normales a priori

Les courbes présentées dans la figure 4.10 présentent :

- l’estimation classique (en trait plein),
- l’estimation bayésienne avec un a priori cohérent et très informatif (en pointillé),

- la valeur théorique (en trait mixte).

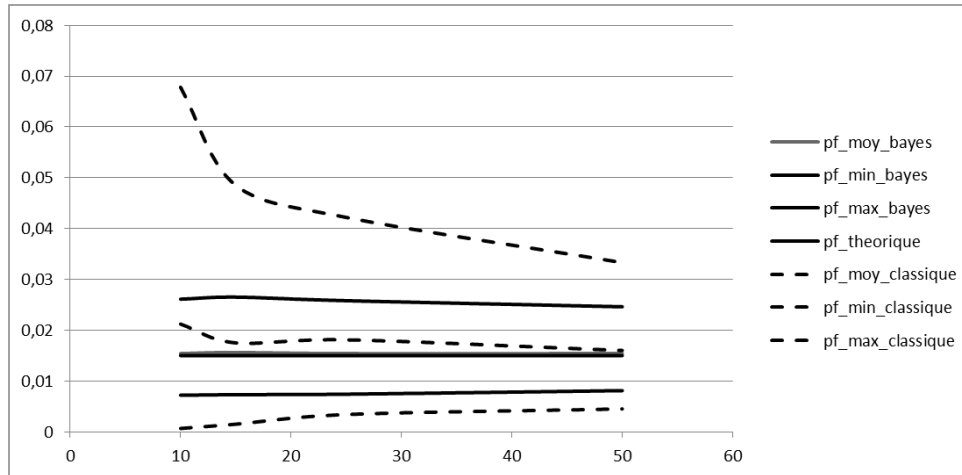


Figure 4.10 – Résultats du Monte Carlo pour les estimations classique et bayésienne de Pf avec les intervalles de confiance à 90%

Nous constatons que pour le cas bayésien avec l'a priori cohérent et très informatif (trait plein) :

- les estimateurs ponctuels bayésiens sont très proches des valeurs théoriques,
- les écart-types des estimateurs sont peu influencés par la taille d'échantillon (cela s'explique par le fait que l'information apportée par les résultats d'essai est faible par rapport à celle contenue dans les a priori),
- les intervalles du paramètres η pour les trois cas, sont relativement proches montrant une certaine insensibilité de ce paramètre à l'apport d'information.

4.4.3 A priori très informatif et incohérent

Pour le troisième cas, une très mauvaise information a priori sur les paramètres a été considérée. Le tableau 4.9 donne l'opinion de l'expert sur les paramètres.

| paramètre | limite inférieure | limite supérieure |
|-----------|-------------------|-------------------|
| C | 400 | 800 |
| β | 1,5 | 2,5 |

Table 4.9 – Prior intervalles

Nous déduisons les paramètres des loi normales a priori par la méthode des moments (tableau 4.10).

| paramètre | moyenne μ | écart-type σ |
|-----------|---------------|---------------------|
| C | 600 | 115,47 |
| β | 2 | 0,29 |

Table 4.10 – Paramètres des lois Normales a priori

Les courbes présentées dans la figures 4.11 présentent :

- l’estimation classique (en pointillé)
- la valeur théorique (trait mixte),
- l’estimation bayésienne, présentée dans ce paragraphe, avec un a priori incohérent et très informatif (trait plein).

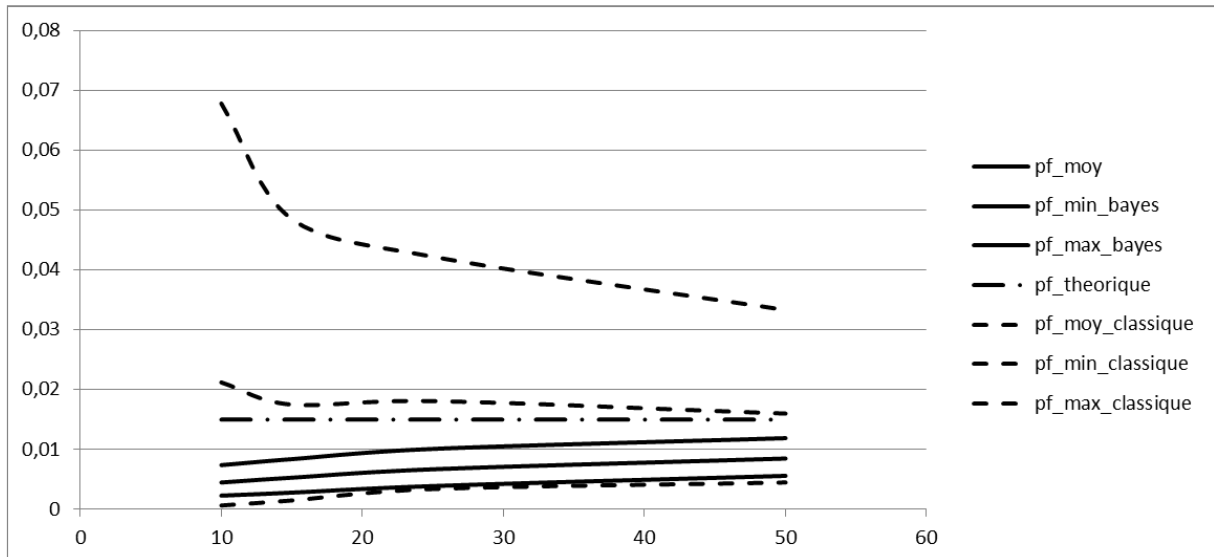


Figure 4.11 – Résultats du Monte Carlo pour les estimations classique et bayésienne de Pf avec les intervalles de confiance à 90%

La forte information a priori s’impose face aux résultats d’essai. Ainsi, à cause de l’écart existant entre les intervalles a priori et les valeurs théoriques, tous les résultats sont décalés par rapport aux vraies valeurs :

- les estimateurs ponctuels pour C et pour β sont sur-évalués,
- l’estimateur ponctuel pour η est sous évalué

Puisque les a priori sont très informatifs, les écart-types sont faibles et peu sensibles aux tailles d’échantillon. Il en découle des intervalles de confiance serrés mais décentrés.

Comme dans ce cas, la connaissance objective donnée par la vraisemblance peut être dégradée par le choix d’un mauvais a priori. L’analyse Bayésienne peut amener à prendre une décision a posteriori erronée. Dans le chapitre 6, nous proposons d’étudier le problème de la compatibilité de l’a priori et les premiers résultats d’essais.

4.5 Conclusion

Dans ce chapitre, nous avons présenté l’estimation bayésienne des modèles SVA paramétriques. La qualité de l’estimation dépendra principalement de la quantité de produits défectueux lors de l’essai. Ainsi, la taille d’échantillon, les niveaux de sollicitation, la durée d’essai, ... sont des paramètres importants qui conditionnent fortement la dynamique de défaillance. Par conséquent, le choix de ces paramètres sont importants et doit faire l’objet d’une attention toute particulière en termes industriel et statistique. Aussi, dans le chapitre suivant nous proposons d’étudier la construction de plans optimaux.

Planification optimisée et robuste d'essais accélérés dans un contexte de qualification

5.1 Introduction

La définition d'un plan d'essai est une étape importante pour garantir une prise de décision (qualification ou non) rigoureuse. La réalisation d'un plan d'essai accéléré est coûteuse en terme de temps, de produits testés, de moyens d'essais et de "monitoring", ... freinant leur mise en place. De plus, lorsque les essais ne permettent pas de précipiter suffisamment de défaillances, dans le cas d'un produit plus fiable que prévu, il est souvent difficile, voir impossible, de prendre une décision quant à la qualification car les intervalles de confiance sont trop larges. D'autre part, un produit qualifié avec un risque important, ou pas, pour lequel le retour d'exploitation (période de garantie, ...) montre un nombre de défaillances supérieur à ce qui était escompté, peut engager financièrement et pénalement l'entreprise.

Ainsi, cela montre l'importance de chercher un compromis optimal entre les coûts d'essai et d'exploitation qui soit peu sensible à la variabilité des paramètres du plan d'essai et à l'incertitude sur la fiabilité réelle du produit.

5.2 Problématique

L'objectif d'un plan d'essai est de juger des performances d'un produit vis à vis d'un cahier des charges et d'une exigence de fiabilité souhaitée. Ainsi, à l'issue de l'essai, il est possible de statuer sur la tenue de l'objectif de fiabilité avec un certain niveau de confiance (niveau de confiance = 1-risque). En effet, les essais étant réalisés sur un nombre limité de produits, il est nécessaire de définir un critère d'acceptation intégrant les intervalles de confiance. La figure 5.1 montre comment peut être défini un critère d'acceptation.

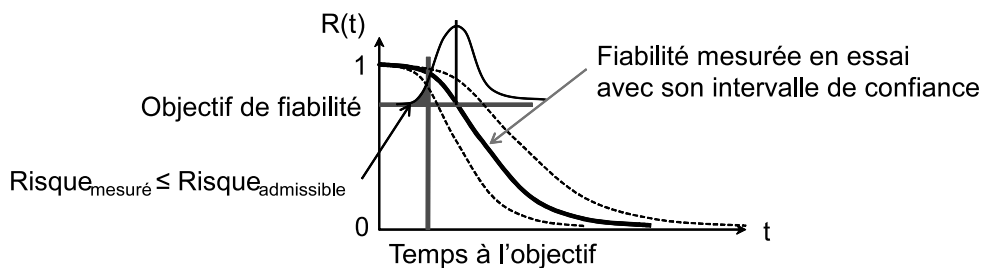


Figure 5.1 – Exemple de critère d'acceptation

Cette exigence est fixée dans les conditions opérationnelles d'utilisation de l'équipement qui peut intégrer un profil de mission plus ou moins complexe. Les différents stress (température, humidité, cyclage thermique, vibration,) peuvent solliciter des modes de défaillance indépendants permettant de décomposer l'essai de qualification en une série d'essais indépendants avec des stress de sévérisation spécifiques.

Ainsi, la construction du plan d'essai consiste à définir la série d'essais accélérés à réaliser et pour chaque essai de proposer un plan optimisé (la meilleure qualité d'estimation pour un coût minimal en termes de taille d'échantillon et de durée d'essai).

5.3 Principe de l’optimisation des plans d’essai

Il existe aujourd’hui plusieurs façons d’appréhender la définition d’un plan d’essai permettant d’estimer tous les paramètres du modèle de vie accélérée. On peut prendre en considération des contraintes industrielles (capacité et performances des moyens d’essai, délai, nombre de produits à tester disponibles,) mais également les aspects statistiques pour rechercher le meilleur plan. Il y a divers types de plan accéléré et en particulier nous pouvons citer les plans : Traditionnel, Traditionnel amélioré; Optimal statistiquement et le meilleur compromis.

- Plans traditionnels ou standards : Essentiellement, ces plans sont définis de manière empirique. Les niveaux de stress bas et haut sont donnés par expérience. De plus, pour réduire le nombre de variables de décision, on fixe quelques paramètres du plan comme : le niveau de stress intermédiaire se trouve au milieu de l’intervalle de stress bas et haut et une répartition égale de l’échantillon total aux différents niveaux. Ces plans fournissent généralement des estimations assez pauvres dans les conditions nominales et sont donc à éviter.
- Plans standards améliorés : Dans ces plans, on fixe quelques paramètres du plan : le niveau de stress intermédiaire se trouve au milieu de l’intervalle de stress bas et haut et une répartition égale de l’échantillon total aux différents niveaux. Toutefois, le niveau bas est optimisé afin de minimiser la variance du percentile de la durée de vie dans les conditions nominales.
- Plans optimal statistiquement : Lorsque l’on a un seul stress d’accélération, les plans utilisent uniquement deux niveaux de stress. Le niveau de stress bas, en respectant une certaine allocation de l’échantillon, est optimisé afin de minimiser la variance du percentile de la durée de vie dans les conditions nominales. Le niveau de stress est fixé préalablement permettant de précipiter beaucoup de défaillances réduisant la variance des estimations. Par contre, ces plans sont très sensibles aux erreurs de spécifications (type de lois, valeurs a priori nécessaires à la construction des plans,). Ainsi, l’optimalité d’un plan peut être impactée par les choix incorrects. De plus, avec seulement deux niveaux de stress il n’est pas possible de vérifier la validité de certaines hypothèses.
- Plans de meilleur compromis : Lorsque l’on a un seul stress d’accélération, les plans utilisent trois (ou plus) niveaux de stress. Le niveau de stress intermédiaire se trouve au milieu de l’intervalle de stress bas et haut. Le niveau de stress bas et l’allocation de l’échantillon total à ce niveau, sont optimisés afin de minimiser la variance du percentile de la durée de vie dans les conditions nominales.

Dans la suite de cette section, nous détaillons le principe de l’optimisation et les hypothèses considérées communément dans la littérature (voir [Yang and Yang, 1998], [Nelson, 1990b], [Meeker and Escobar, 1998]...).

5.3.1 Hypothèses générales

Il est possible de formuler le problème de la recherche d’un plan optimal en posant les hypothèses, présentées ci-dessous, permettant de traiter la plupart des problèmes de plans d’essais accélérés.

1. Le logarithme du temps à défaillance de chaque équipement suit une distribution « position et échelle » telle que :

$$prob(Y \leq y) = \Phi\left(\frac{y-\mu}{\sigma}\right)$$

Où μ et σ sont les paramètres de position et d'échelle de la distribution et $\Phi(\cdot)$ est la forme standard d'une distribution « position-échelle ».

2. Les temps de défaillances de tous les équipements en test, à tous les niveaux de stress, sont statistiquement indépendants.
3. Le paramètre de position μ est une fonction linéaire du stress z_1 . Plus particulièrement, nous traiterons : $\mu = \mu(z_1) = \gamma_0 + \gamma_1 \cdot z_1$
4. Le paramètre d'échelle σ ne dépend pas des niveaux de stress. Tous les équipements sont testés jusqu'à un temps d'essai fixé préalablement.
5. Deux des principales lois utilisées dans les essais accélérés sont les lois de Weibull et LogNormale. La loi de Weibull est donnée par : $Y \propto DVE(\mu(z_1) = \gamma_0 + \gamma_1 \cdot z_1, \sigma)$
Où DEV représente la distribution des valeurs extrêmes. La loi logNormale est donnée par : $Y \propto Normal(\mu(z_1) = \gamma_0 + \gamma_1 \cdot z_1, \sigma)$. C'est le logarithme de la durée de vie Y qui est décrit par une distribution DEV et Normale de paramètre de position $\mu(z)$, défini comme une fonction linéaire de stress z, et d'un paramètre d'échelle constant ?.

5.3.2 Critère de planification et formulation du problème

Il est possible de standardiser le stress, sans perdre la généralité, par :

$$\xi = \frac{x - x_N}{x_H - x_N} \tag{5.1}$$

Où

- x_N est le niveau nominal du stress correspondant aux conditions opérationnelles de l'équipement.
- x_H est le niveau le plus élevé du stress.

Les valeurs de x , x_D et x_H sont associées aux valeurs de stress ou à leurs transformations (par exemple l'inverse de la température pour la loi d'Arrhenius ou la transformation logarithmique pour la loi puissance inverse). La valeur la plus élevée du stress doit être choisie de telle sorte que l'on conserve le même mécanisme de défaillance et par conséquent les mêmes hypothèses de lois de fiabilité et d'accélération. Elle peut être choisie en considérant les aspects technologiques tout en sachant que ce niveau permettra de précipiter rapidement des défaillances. Finalement, nous obtenons $\xi=0$ pour le niveau nominal et $\xi=1$ pour le niveau le plus élevé. La plupart du temps, les essais accélérés sont construits pour estimer une valeur particulière de percentile T_p (probabilité de défaillance p), de la distribution des durées de vie en conditions nominales. Le critère naturel est d'optimiser $V(\hat{T}_p)$ ou $V(\hat{Y}_p)$ avec $Y_p = \ln(T_p)$. $V(\hat{Y}_p)$ mesure la précision de l'estimation du quantile p. Un « bon » plan doit permettre d'obtenir une valeur faible, si n'est pas le minimum, de $V(\hat{Y}_p)$. Pour le problème de minimisation, les variables de décision sont ξ_i (les niveaux de stress standardisés) et π_i (les pourcentages de la taille d'échantillon total alloués aux différents niveaux de stress). Le problème d'optimisation peut être formulé de la manière suivante :

$$\left| \begin{array}{l} \text{Minimiser } V(\hat{Y}_p) = f(\xi_i, \pi_i) \\ \text{sous contrainte : } 0 \leq \pi_i \leq 1, i = 1, 2, \dots, n \\ \text{où } \sum_{i=1}^n \pi_i = 1 \end{array} \right. \tag{5.2}$$

La recherche d'un plan optimal nécessite de recourir à des algorithmes pour minimiser $V(\hat{Y}_p)$. La construction de plans doit être un compromis entre efficacité et précision d'extrapolation. Plus il y a de défaillances, meilleure est l'efficacité d'estimation nécessitant de hauts niveaux de stress mais imposant une extrapolation importante pour passer aux conditions nominales. Choisir le « meilleur » plan consiste donc à rechercher le meilleur compromis entre efficacité et extrapolation. Les plans d'essai avec plusieurs niveaux de stress sont plus robustes que les plans avec peu de niveaux car ils ne permettent pas de tester la validité du modèle de vie accélérée. Toutefois, ces derniers plans sont plus simples à mettre en place.

5.3.3 Plan d'essai pour un seul stress

Ce paragraphe présente plusieurs plans couramment utilisés quand seulement un stress est appliqué sur les équipements en essai. Pour construire les plans, il est nécessaire de fournir quelques informations :

- Le stress nominal x_N et le stress le plus élevé x_H .
- La durée d'essai (temps de censure) η_i par niveau de stress.
- La valeur du paramètre d'échelle σ .

Ces informations permettent de formuler $V(\hat{Y}_p)$.

Formulation de $V(\hat{Y}_p)$:

Nous avons $\hat{Y}_p = \hat{\gamma}_0 + \hat{\gamma}_1 \cdot x_N + z_p \cdot \sigma$

Où z_p est le percentile estimé par $[z - p = \log[-\log(1-p)]]$ pour le cas d'une loi de valeurs extrêmes SEV (Weibull) ou par $[z_p = \Phi_{nor}^{-1}(p)]$ dans le cas de la loi normale (LogNormale) Alors la variance asymptotique $V(\hat{Y}_p)$ peut être estimée par

$$V(\hat{Y}_p) = [1 \ x_N \ z_p] \Sigma \begin{bmatrix} 1 \\ x_N \\ z_p \end{bmatrix}$$

Où la matrice de variance-covariance Σ est donnée par

$$\Sigma = \begin{bmatrix} Var(\hat{\gamma}_0) & Cov(\hat{\gamma}_0, \hat{\gamma}_1) & Cov(\hat{\gamma}_0, \hat{\sigma}) \\ Cov(\hat{\gamma}_0, \hat{\gamma}_1) & Var(\hat{\gamma}_1) & Cov(\hat{\gamma}_1, \hat{\sigma}) \\ Cov(\hat{\gamma}_0, \hat{\sigma}) & Cov(\hat{\gamma}_1, \hat{\sigma}) & Var(\hat{\sigma}) \end{bmatrix}$$

Et la matrice de Fisher par $I = \sum_{i=1}^n F_i$

$$\text{avec } \Sigma = \begin{bmatrix} A(\varsigma_i) & x_i A(\varsigma_i) & B(\varsigma_i) \\ x_i A(\varsigma_i) & x_i^2 A(\varsigma_i) & x_i B(\varsigma_i) \\ B(\varsigma_i) & x_i B(\varsigma_i) & C(\varsigma_i) \end{bmatrix}$$

où $\varsigma_i = (\eta_i - \gamma_0 - \gamma_1 \cdot x_i) / \sigma$ avec $i = (1, 2, 3)$ représentant respectivement les niveaux de stress x_B , x_M et x_H .

Les termes $A(\cdot)$, $B(\cdot)$ et $C(\cdot)$ sont déterminés par : $A(\varsigma_i) = \Phi(\eta_i - \gamma_0 - \gamma_1 \cdot x_i) / \sigma$

$$B(\varsigma_i) = \int_0^{e^{\varsigma_i}} \ln(w) w e^{-w} dw + (1 - \Phi(\varsigma_i)) \varsigma_i^2 e^{\varsigma_i}$$

$$C(\varsigma_i) = \left\{ \Phi(\varsigma_i) + \int_0^{e^{\varsigma_i}} \ln(w) w e^{-w} dw + (1 - \Phi(\varsigma_i)) \varsigma_i^2 e^{\varsigma_i} \right\}$$

Dans les articles de Escobar et Meeker ([Meeker and Escobar, 1986a], [Meeker and Escobar, 1986b] et [?]), on trouve les algorithmes permettant d'estimer les termes $A(\cdot)$, $B(\cdot)$ et $C(\cdot)$.

Plusieurs plans optimaux sont couramment proposés dans la littérature : **Plan optimal statistiquement à deux niveaux de stress** Le plan optimal statistiquement à deux niveaux de stress est le plus important des plans. Pour ce plan, le niveau de stress le plus élevé x_H et celui au niveau nominal x_D sont prédéterminés. Le niveau de stress bas d'essai x_B , et la

taille d'échantillon allouée à ce niveau x_B , sont calculés tels qu'ils minimisent $V(\hat{Y}_p)$. **Meilleur Plan standard à trois niveaux de stress** Dans ce plan, 3 niveaux de stress sont utilisés. Considérons ξ_B, ξ_M, ξ_H les 3 niveaux de stress standardisés avec le niveau moyen ξ_M au milieu des niveaux bas ξ_B et haut ξ_H :

Une répartition égale de la taille d'échantillon total est considérée à chaque niveau de test ($\pi_B = \pi_M = \pi_H = 1/3$). Ainsi, le plan d'essai est $(\zeta_B, \zeta_M, \zeta_H, \pi_B, \pi_M, \pi_H) = (\zeta_B, (\zeta_B+1)/2, 1, 1/3, 1/3, 1/3)$ où seule la variable ζ_B est à estimer telle qu'elle minimise $V(\hat{Y}_p)$.

Meilleur compromis de plan à trois niveaux de stress

Dans ce plan, 3 niveaux de stress sont utilisés tels que $(\zeta_B, (\zeta_B+1)/2, 1)$. La proportion de la taille d'échantillon total au niveau moyen π_M est égale à 0.2 (ou 0.1) qui donne de bons résultats. Ainsi, le plan d'essai est $(\zeta_B, \zeta_M, \zeta_H, \pi_B, \pi_M, \pi_H) = (\zeta_B, (\zeta_B+1)/2, 1, \pi_B, \pi_M, 1-\pi_B-\pi_M)$ où seules les variables ζ_B et π_B sont à estimer telles qu'elles minimisent .

5.4 Analyse des méthodes existantes de construction de plans optimaux

Pour ce faire, nous allons considérer un exemple. Un essai accéléré doit être conduit à plusieurs niveaux de tension pour estimer le quantile 10 % de la durée de vie d’un condensateur chimique à la tension nominale 15kV. L’essai ne devra pas dépasser une durée de 672 heures. Le nombre total d’unités disponible pour l’essai est de 100 condensateurs. Pour éviter d’introduire de nouveaux modes de défaillance que ceux observés lors des conditions opérationnelles, il est décidé, d’après les avis d’expert, de ne pas dépasser la tension de 38kV. De plus, on fait l’hypothèse que la distribution des durées de vie est de type Weibull et que la loi d’accélération est un modèle puissance inverse. Le tableau 5.1 suivant donne le résultat pour les 3 plans présentés précédemment.

| | | Plan optimal statistiquement à deux niveaux de stress | Meilleur Plan standard à trois niveaux de stress | Meilleur compromis de plan à trois niveaux de stress |
|--|-------------|---|--|--|
| Stress standardisé | ζ_B | 0.0535 | 0.3933 | 0.4943 |
| | ζ_M | - | 0.6967 | 0.7472 |
| | ζ_H | 1 | 1 | 1 |
| Proportion de l’échantillon allouée par niveau de stress | π_B | 0.7665 | 0.3333 | 0.5840 |
| | π_M | - | 0.3333 | 0.2 |
| | π_H | 0.2335 | 0.3333 | 0.2160 |
| Probabilité de défaillance | p_B | 0.2980 | 0.1180 | 0.2254 |
| | p_M | - | 0.6533 | 0.7793 |
| | p_H | 0.9999 | 0.9999 | 0.9999 |
| Variance standardisée | $V_{Y10\%}$ | 27.47 | 36.42 | 31.90 |

Table 5.1 – Résultats pour plusieurs types de plans

Cet exemple montre que le plan optimal statistiquement à deux niveaux de stress est celui qui permet d’obtenir la variance la plus faible 27.47. Par contre, ce plan est très sensible aux erreurs de spécifications (type de lois, valeurs a priori nécessaires à la construction des plans,). Ainsi, l’optimalité du plan peut être impactée par les choix incorrects. De plus, avec seulement deux niveaux de stress il n’est pas possible de vérifier la validité de certaines hypothèses. Le plan meilleur compromis à trois niveaux de stress permet de lever ces difficultés avec un variance un tout plus élevée 31.90.

Dans le cas où la fiabilité réelle est bien meilleure, ou plus mauvaise, l’analyse des résultats d’essai conduira à une estimation dégradée (par exemple, aucune défaillance observée aboutira à aucune estimation).

Il est donc important, de proposer une méthode de construction de plan permettant de palier

à ces différentes difficultés.

5.5 Proposition d'un plan d'essai optimal et robuste

L'objectif de la méthode proposée est de définir un plan d'essais accélérés optimisé en considérant une approche économique. La fonction objectif est définie par deux termes :

- les coûts associés à l'activité d'essai
- les coûts reliés à l'exploitation des équipements

Dans de nombreux travaux, les plans d'essais optimaux sont définis en considérant les connaissances a priori sur les paramètres de fiabilité (choix de fonction de fiabilité, valeurs des paramètres d'échelle et de forme, ...) et de modèle d'accélération (choix de loi d'accélération, valeurs des paramètres de lois, ...) pour évaluer les probabilités de défaillance aux différents niveaux sévérés.

Cette connaissance a priori n'est pas intégrée dans l'inférence statistique lors du processus d'optimisation (minimisation de la variance de la métrique de fiabilité) et de l'analyse des résultats d'essai (estimation des paramètres du modèle SVA). Pourtant, cette information peut être utilisée dans une approche bayésienne pour optimiser le plan d'essai. De plus, elle contient également l'incertitude sur la fiabilité réelle du nouvel équipement.

Aussi, nous proposons une méthodologie qui consiste à définir un plan d'essai optimal en considérant une fonction objectif basée sur une approche économique, l'inférence bayésienne pour optimiser le plan et la prise en compte de l'incertitude sur les paramètres du modèle SVA pour obtenir un plan robuste.

La figure 5.2 présente les différentes étapes de méthodologie.

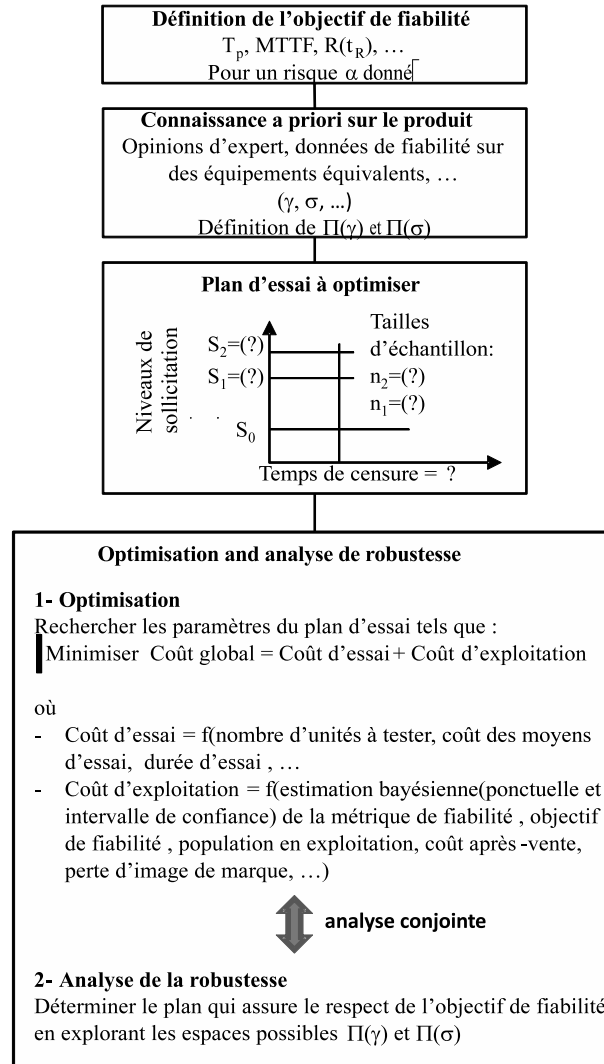


Figure 5.2 – Principe de la méthode proposée

5.5.1 Définition de l'objectif de fiabilité

Le plan de test est construit pour démontrer un objectif de fiabilité. Différentes métriques sont utilisées pour caractériser la fiabilité d'un produit : MTTF, B10, probabilité de défaillance sur la période de garantie, ... La qualification consiste à évaluer le risque de ne pas atteindre l'objectif de fiabilité en considérant la borne de l'intervalle de confiance unilatéral (voir figure 5.3).

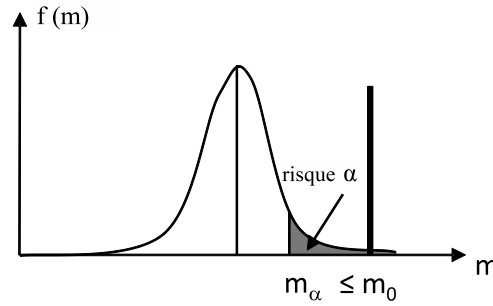


Figure 5.3 – Définition du critère de qualification

Ainsi, la métrique de fiabilité m_α est définie par une fonction $m(\gamma, \sigma)$ qui dépend des paramètres du modèle SVA.

La qualification du produit est validée lorsque la métrique mesurée vérifie :

$$Prob(m_\alpha \geq m_0) \leq \alpha \quad (5.3)$$

Avec m_0 l'objectif de fiabilité à tenir, α le risque fixé et m_α la borne unilatérale.

Remarque : Il est possible que certaine métrique se formule par une inégalité inverse. Toutefois, le problème général se formule de la même manière et dans la suite de la thèse nous continuerons avec cette formulation.

La borne unilatérale m_α est estimée par :

$$\begin{aligned} - m_\alpha &= \hat{m} e^{\frac{-u_{1-\alpha} V(m)}{\hat{m}}} \text{ pour } m \in [0, +\infty[\\ - m_\alpha &= \left[\frac{\hat{m}}{\hat{m} + (1-\hat{m}) e^{\frac{-u_{1-\alpha} \sqrt{V(m)}}{\hat{m}(1-\hat{m})}} \right] \hat{m} e^{\frac{-u_{1-\alpha} V(m)}{\hat{m}}} \text{ pour } m \in [0, 1] \end{aligned}$$

avec

$$\begin{aligned} - \hat{m} &= m(\hat{\gamma}, \hat{\sigma}) \\ - V(m) &\approx \sum_{l=0}^m \left(\frac{\partial m(\gamma, \sigma)}{\partial \gamma_l} \right)_{\hat{\gamma}, \hat{\sigma}}^2 V(\gamma_l) + \left(\frac{\partial m(\gamma, \sigma)}{\partial \sigma} \right)_{\hat{\gamma}, \hat{\sigma}}^2 V(\sigma) \end{aligned}$$

Les variances $V(\gamma_l)$ et $V(\sigma)$ sont obtenues par la méthode MAP.

5.5.2 Plan d'essai à optimiser

Selon le degré de connaissance sur la fiabilité du produit et la loi d'accélération, nous pouvons être amenés à sélectionner a priori le type de plan (statistiquement optimal, meilleur compromis...). Une fois ce choix réalisé, on définit les paramètres du plan à rechercher par optimisation :

- les niveaux de stress (S_i),

- les temps de censure à chaque niveau τ_i ,
- les tailles d'échantillon allouées (n_i) à chaque niveau,
- ...

Certains des paramètres peuvent être fixés pour prendre en compte, par exemple, des considérations technologiques (par exemple la capacité des moyens d'essai), et/ou industrielles (par exemple le délai de qualification) etc.

Ainsi, le plan à optimiser sera désigné par $P(\omega)$ avec ω les paramètres à optimiser.

5.5.3 Connaissance a priori sur le produit

La connaissance a priori sur le nouveau produit (γ, σ, \dots) est obtenue des experts, des données de fiabilité issues d'équipements équivalents, des recueils de données de fiabilité, ... On considère l'hypothèse d'indépendance entre les variables et ainsi la distribution a priori est définie par :

$$\pi(\gamma, \sigma) = \pi(\sigma) \left[\prod_{l=0}^m \pi_{\gamma_l}(\gamma_l) \right] \quad (5.4)$$

Le choix de la forme de $\pi(\cdot)$ dépend du degré de connaissance sur les paramètres (γ_l et σ) et/ou de l'analyse de construction d'une distribution pseudo conjuguée bayésienne selon la procédure présentée dans le paragraphe 4.3.2.

5.5.4 Définition de la fonction coût

On définit la fonction coût $C(\cdot)$ en considérant une formulation économique. En effet, la définition d'un plan d'essai de qualification consiste à statuer sur la fiabilité d'un équipement vis à vis de l'objectif contractuel au moindre coût et pour un risque acceptable de se tromper dans la prise de décision. Ainsi, deux dimensions antagonistes interviennent dans le coût de la qualification, celle :

- associée à l'activité d'essai (Moyens d'essai, Développement du monitoring, Échantillon de produits à tester, ...),
- reliée à la non tenue de l'objectif de fiabilité en exploitation impliquant une prise en charge des coûts de non qualité (Rechange, réparation, pénalités, perte de marché, ...).

Lorsque l'on diminue le risque de se tromper dans la prise de décision quant à la qualification (diminution du coût d'exploitation), cela implique de réduire la variance d'estimation de la métrique de fiabilité. Par conséquent, le coût d'essai augmente notablement puisqu'il sera nécessaire de tester un échantillon de produits important et suffisamment longtemps pour précipiter des défaillances. Inversement, lorsque que l'on réduit le coût d'essai on augmente le risque de ne pas tenir l'objectif de fiabilité et d'augmenter le coût d'exploitation. Aussi, nous proposons de définir une fonction coût intégrant ces deux dimensions :

$$C(\omega, \vartheta, m_\alpha(\gamma, \sigma), m_0) = C_{essai}(\omega, \vartheta, m_\alpha(\gamma, \sigma), m_0) + C_{exploitation}(\omega, \vartheta, m_\alpha(\gamma, \sigma), m_0) \quad (5.5)$$

avec

- ω les paramètres du plan à optimiser,
- ϑ les paramètres d'estimation des coûts,
- $m_\alpha(\gamma, \sigma)$ l'estimateur de la métrique au niveau de risque α déduit du modèle SVA de paramètres γ et σ ,

– m_0 la valeur de la métrique contractuelle.

La figure 5.4 illustre ces deux dimensions de la fonction coût en indiquant le paramètre optimal $\tilde{\omega}_i$ du plan $P(\omega)$.

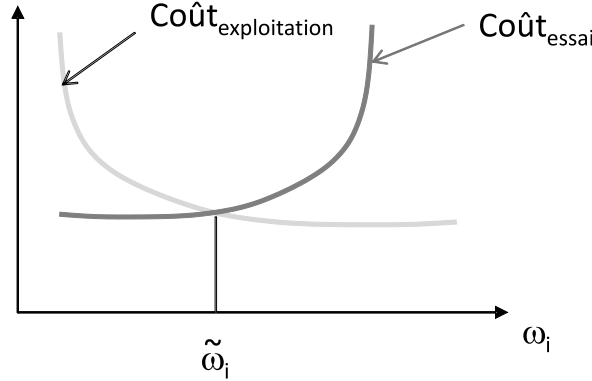


Figure 5.4 – Définition des coûts et paramètre de plan optimal

5.5.5 Formulation du problème d’optimisation

Finalement, le problème d’optimisation se formule de la manière suivante :
Rechercher les paramètres du plan ω du plan choisi $P(\omega)$ tels que

$$\left| \begin{array}{l} \text{Minimiser } C(\omega, \vartheta, m_\alpha(\gamma, \sigma), m_0) \\ \text{sous contraintes :} \\ - g_i(\omega_i) \geq 0, i = 1, \dots, r \\ - \gamma_l \in \pi_l(\gamma_l) l = 1, \dots, m \\ - \sigma \in \pi(\sigma) \end{array} \right. \quad (5.6)$$

La fonction $g_i(\omega_i)$ représente la fonction contrainte sur la contrainte ω_i comme par exemple $\sum n_i = n$. Cette formulation permet de fixer certains paramètres de plan d’essai (comme par exemple le niveau intermédiaire de stress fixé au milieu de l’intervalle défini par les valeurs min et max) et de prendre en compte le domaine possible des paramètres de modèle SVA pour obtenir un plan optimal robuste vis à vis de la fiabilité réelle du produit testé.

5.6 Méthodes d’optimisation d’un plan d’essai

La recherche d’un optimum à la formulation du coût telle que présenté dans dans le paragraphe 5.5.5 n’est pas résoluble analytiquement. L’emploi de méthodes numériques s’impose pour trouver le plan de test optimum du point de vue du coût. Plusieurs algorithmes d’optimisation peuvent être appliqués : surface de réponse, algorithme du gradient stochastique, recuit simulé... Dans la suite nous proposons deux approches d’optimisation :

- en utilisant une surface de réponse,
- par les algorithmes génétiques.

5.6.1 Optimisation à l'aide d'une surface de réponse

La méthode de surface de réponse (RSM), introduite par Box et Wilson en 1951, est un ensemble de techniques développées pour rechercher des paramètres de conception $x = x_1, \dots, x_k$ (par exemple les paramètres d'un plan) optimaux qui optimisent une réponse cible Y (par exemple la minimisation de la fonction coût). Pour ce faire, la méthode de surface de réponse consiste à construire une expression approchée \tilde{f} de la fonction objective $Y = f(x)$ à partir d'un nombre limité d'évaluations de la fonction réelle. L'ensemble des points x est choisi à partir d'un plan d'expérience, et la fonction $f(x)$ est construite en utilisant une méthode d'approximation. Plusieurs modèles mathématiques peuvent être utilisés :

- modèle du premier ordre (sans interaction entre les variables) :

$$\tilde{f}(x) = b_0 + \sum_{i=1}^k b_i \cdot x_i + \epsilon$$

- modèle du deuxième ordre (avec interaction entre les variables) :

$$\tilde{f}(x) = b_0 + \sum_{i=1}^k b_i \cdot x_i + \sum_{j=2}^k \sum_{i=1}^{j-1} b_{ij} x_i x_j + \sum_{i=1}^k b_i \cdot x_i^2 + \epsilon$$

Pour la minimisation de $f(x)$, on peut utiliser des méthodes (Méthode des gradients conjugués, BFGS, ...) faisant un grand nombre d'appels à la fonction $f(x)$, car l'évaluation fréquente de la fonction approchée ne pose plus de problèmes.

Ainsi, dans notre cas la méthode consiste à construire une surface de réponse sur l'espace des solutions permettant de définir le coût global $C(\omega, \vartheta, m_\alpha(\gamma, \sigma), m_0)$ en fonction des paramètres du plan ω . La construction se fera pas simulation de Monte Carlo assurant une mise en oeuvre simple mais qui peut être coûteuse en temps de calcul. Pour commencer, il est nécessaire de définir un certain nombre de paramètres caractérisant la qualification à traiter :

- le plan d'essai $P(\omega)$ à optimiser et les paramètres associés $\omega = \omega_1, \dots, \omega_r$,
- l'intervalle possible de solution pour chaque paramètre $\omega_i \in [\omega_i^{min}, \omega_i^{max}]$ avec $i = 1, \dots, r$,
- le modèle SVA $R_S(t) = R\left(\frac{\ln(t) - \gamma^T \cdot Z^0}{\sigma}\right)$ caractérisant la fiabilité en fonction des sollicitations de sévérisation S et des paramètres $\gamma = \gamma_1, \dots, \gamma_m$ et σ ,
- les distributions a priori $\Pi_{\gamma_i}(\gamma_i)$ et $\Pi_\sigma(\sigma)$ pour les paramètres γ_i et σ ,
- l'objectif de fiabilité contractuelle m_0 , fonction coût $C(\omega, \vartheta, m(\gamma, \sigma), m_0)$ et les paramètres de coût associés ϑ ,

Ensuite, on construit la surface de réponse en discrétisant l'espace des solutions $[\omega_{min}, \omega_{max}]$ et pour chaque point $\check{\omega}$ de cette espace on détermine la fonction coût $C(\check{\omega}, \vartheta, m_\alpha(\gamma, \sigma), m_0)$ par simulation de Monte Carlo afin de prendre en compte les effets d'échantillonnage et les incertitudes sur la fiabilité réelle et la loi d'accélération (caractérisée par $\Pi_{\gamma_i}(\gamma_i)$ et $\Pi_\sigma(\sigma)$). Les étapes de construction sont les suivantes :

1. discrétisation de l'espace des solutions

$$\check{\omega} = \left\{ \left\{ \omega_1^{min}, \dots, \omega_1^i, \dots, \omega_1^{max} \right\} \dots \left\{ \omega_j^{min}, \dots, \omega_j^i, \dots, \omega_j^{max} \right\}, \dots \left\{ \omega_r^{min}, \dots, \omega_r^i, \dots, \omega_r^{max} \right\} \right\}$$

2. génération aléatoire des paramètres $\tilde{\gamma}$ et $\tilde{\sigma}$ du modèle SVA $R_S(t) = R\left(\frac{\ln(t) - \gamma^T \cdot Z^0}{\sigma}\right)$ à partir des lois a priori $\Pi_\gamma(\gamma)$ et $\Pi_\sigma(\sigma)$ en respectant les paramètres $\check{\omega}$
3. génération aléatoire des résultats d'essai respectant le modèle SVA $R_S(t) = R\left(\frac{\ln(t) - \gamma^T \cdot Z^0}{\sigma}\right)$, le plan $P(\check{\omega})$ et ses paramètres $\check{\omega}$,
4. estimation MAP des paramètres de modèle SVA $(\hat{\gamma}, \hat{\sigma})$ à partir des résultats d'essai simulés et détermination de la métrique $m_\alpha = m(\hat{\gamma}, \hat{\sigma})$

5. répétition k fois de l'étape 2 à 4 afin d'obtenir un échantillon de métriques $m_\alpha = m_{\alpha, \dots, m_\alpha}^1$ (cette répétition permet de balayer le domaine possible de la fiabilité réelle et ainsi de définir un plan robuste),
6. estimation du percentile $m_\alpha^{\alpha'}$ de la métrique au risque α'
7. évaluation de la fonction coût global $C(\hat{\omega}, \vartheta, m_\alpha^{\alpha'}, m_0)$
8. répétition des étapes 1 à 5 pour chaque point de l'espace discrétisé des solutions pour construire l'ensemble de la surface de réponses.

Pour finir, l'hypersurface du coût global est approximée par un polynôme quadratique pour lisser les effets de simulation de Monte Carlo. Le plan d'essai optimum est obtenu par une méthode d'optimisation classique en minimisant le coût défini par le polynôme.

5.6.2 Exemple numérique pour un plan de qualification de roulements à billes

Dans ce paragraphe, la méthodologie est appliquée à la définition d'un plan de qualification de roulements à billes. Pour cela, on reprend les éléments de simulation détaillés dans les chapitres précédents.

5.6.2.1 Définition de la métrique de qualification

Le plan de test est construit pour démontrer une probabilité de défaillance p sur la durée de garantie.

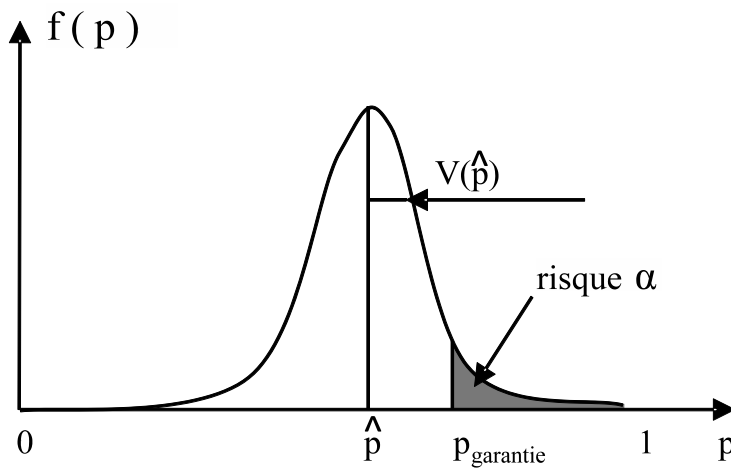


Figure 5.5 – Fonction de densité de l'estimation de la probabilité de défaillance

L'estimateur ponctuel est défini par :

$$\hat{p} = 1 - R_{X_0}(T_{garantie}) = R\left(\left(\ln(T_{garantie}) - \hat{\gamma}^T Z_0\right) / \hat{\sigma}\right) \quad (5.7)$$

Et la variance

$$V(\hat{p}) \approx \sum_{l=0}^m \left(\frac{\partial(1 - R_{X_0}(T_{garantie}))}{\partial \gamma_l}\right)_{\hat{\gamma}_l, \hat{\sigma}}^2 V(\gamma_l) + \left(\frac{\partial(1 - R(t))}{\partial \gamma_l}\right)_{\hat{\gamma}_l, \hat{\sigma}}^2 V(\sigma) \quad (5.8)$$

Les paramètres $\hat{\gamma}_1$, $\hat{\sigma}$, $V(\gamma_l)$ et $V(\sigma)$ seront obtenus par simulation de Monte Carlo définie après. La probabilité de défaillance sur la période de garantie est estimée par la borne maximale de l'intervalle unilatéral pour le risque α (voir figure 5.5). Pour résoudre des problèmes numérique d'estimation, la distribution Beta $f(p)$ est considérée pour caractériser la distribution de p . Ses paramètres sont estimés par la méthode des moments :

$$\alpha_0 = \frac{\hat{p}^2}{V(\hat{p})}(1 - \hat{p}) - \hat{p}$$

$$\beta_0 = \frac{\hat{p}}{V(\hat{p})}(1 - \hat{p})^2 + \hat{p} - 1$$

La probabilité de défaillance est déduite de :

$$p_{garantie} = Beta^{-1}(1 - \alpha | \alpha_0, \beta_0) \quad (5.9)$$

Pour finir, la qualification est validée lorsque $p_{garantie} \leq p_{objectif}$ avec $p_{objectif}$ la probabilité contractuelle.

5.6.2.2 Définition du plan de test

Pour cet exemple, un cas simple de plan d'essai (avec deux niveaux de stress constants) est traité avec les paramètres suivants :

- Temps de censure unique τ fixé
- Taille d'échantillon n donnée
- Stress maxi S_2 fixé
- Stress intermédiaire S_1 inconnu
- Proportion de la taille d'échantillon p_1 allouée au niveau de Stress S_1 inconnue ($n_1 = \text{ent}(p_1 \cdot n)$ and $n_2 = n - n_1$)

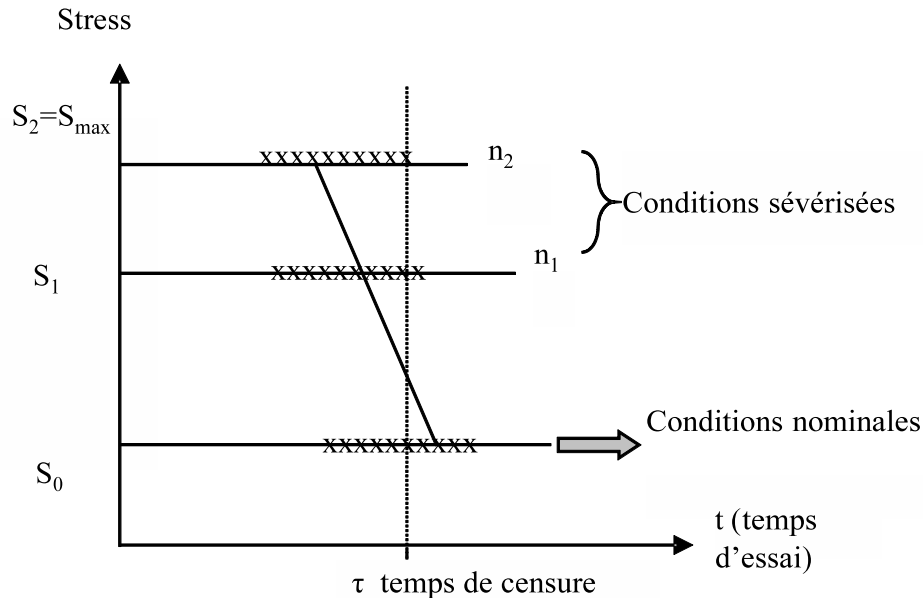


Figure 5.6 – Plan d'essai à optimiser

Les paramètres du plan à rechercher par optimisation sont :

- La proportion de la taille d'échantillon p_1 allouée au niveau de Stress S_1 .
- Le stress intermédiaire S_1 .

5.6.2.3 Modèle SVA pour les roulements à billes

Le modèle SVA Weibull-Puissance Inverse est celui qui sera utilisé pour simuler les instants de défaillance pour les roulements à billes. Il est défini par :

$$R(u, S) = e^{-\left(\frac{u}{\left(\frac{C}{S}\right)^p \frac{60 \cdot N}{10^6} \left(\frac{1}{0.105}\right)^{1.5}}\right)^{1.5}} \tag{5.10}$$

avec

- C est la charge dynamique de base
- S est la charge radiale équivalente
- $p = 3$ pour les roulements à billes et $p = 10/3$ pour les roulements à rouleaux
- N la vitesse de rotation.

Le paramètre de forme est fixé à 1.5. Le plan d'essai de qualification est construit pour des roulements à billes dont les caractéristiques sont données dans le tableau 5.2.

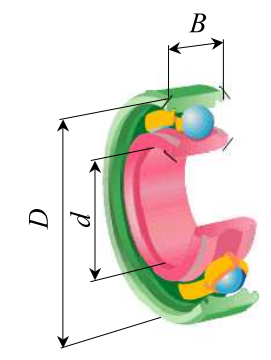
| | | | | | | |
|--|--|--------------------------------|----------------------------------|----------------------|--|---|
|  | <table border="1"> <tr> <td>$D = 32$ mm : outside diameter</td> </tr> <tr> <td>$d = 15$ mm : diamètre intérieur</td> </tr> <tr> <td>$B = 8$ mm : largeur</td> </tr> <tr> <td>$C = 400$ daN : charge dynamique de base</td> </tr> <tr> <td>$N = 1500$ tr/min : vitesse de rotation</td> </tr> </table> | $D = 32$ mm : outside diameter | $d = 15$ mm : diamètre intérieur | $B = 8$ mm : largeur | $C = 400$ daN : charge dynamique de base | $N = 1500$ tr/min : vitesse de rotation |
| $D = 32$ mm : outside diameter | | | | | | |
| $d = 15$ mm : diamètre intérieur | | | | | | |
| $B = 8$ mm : largeur | | | | | | |
| $C = 400$ daN : charge dynamique de base | | | | | | |
| $N = 1500$ tr/min : vitesse de rotation | | | | | | |

Table 5.2 – Caractéristiques du roulement

5.6.2.4 Connaissance a priori sur le produit

La connaissance a priori disponible sur le nouveau roulement à billes (γ, σ, \dots) est obtenue des experts qui donnent des intervalles sur les deux paramètres inconnus (C, ν). On considère l'hypothèse d'indépendance entre les variables et ainsi la distribution a priori est définie par :

$$\pi(C, \nu) = \pi(C) \cdot \pi(\sigma) \tag{5.11}$$

Pour des questions de simplicité, nous avons choisi des lois normales pour définir les distributions a priori.

5.6.2.5 Fonction objectif

L'objectif de l'optimisation du plan d'essai accéléré est de minimiser le coût global (intégrant les coûts d'essai et d'exploitation), qui est défini par :

$$Coût_{global} = Coût_{essai} + Coût_{exploitation} \quad (5.12)$$

Le coût d'essai intègre :

- le coût des produits testés : $n \cdot \text{coût unité} + \text{coût fixe d'essai}$,
- le coût fixe : coût fixe d'essai
- le coût d'essai associé à la durée de l'essai : $\text{coût par heure d'essai} \cdot \text{Durée}$
- le coût lié au nombre de séries d'essais à mettre uvre : $(n_1 / \text{nombre maxi d'unités par lot d'essai})^{\gamma_{essai}} + (n_2 / \text{nombre maxi d'unités par lot d'essai})^{\gamma_{essai}}$

Le paramètre γ_{essai} permet de matérialiser une augmentation non linéaire du coût associé au nombre de séries d'essai (immobilisation de plusieurs moyens d'essai, nécessité d'avoir plusieurs monitoring, ...).

Le coût d'exploitation permet de prendre en compte :

- le coût fixe de prise en charge de la garantie : $\text{coût fixe garantie}$,
- le coût de prise en charge des défaillances lors de la période de garantie : $(p_{exploitation} - p_{cible}) \cdot \text{taille de la population en exploitation} \cdot \text{coût de l'unité}$,
- le coût de perte d'image de marque de l'entreprise lorsque la proportion de défaillance est supérieure à l'objectif : $\text{Coût de la perte d'image} \cdot (p_{exploitation} - p_{cible})^{\gamma_{exploitation}}$

Le paramètre $\gamma_{garantie}$ permet de matérialiser une augmentation non linéaire du coût de perte d'image de marque, de part de marché, ...

5.6.2.6 Procédure d'optimisation

La procédure d'optimisation est simplifiée en définissant l'espace des solutions sur le domaine possible des variables ($p_1 \in]0, 1[$, $S_1 \in]S_0, S_2[$). Le domaine possible est discrétisé selon la grille suivante:

$$p_1 = \{p_{1min}, \dots, p_{1i}, \dots, p_{1max}\} \text{ avec } i = 0, m_1$$

$$S_1 = \{S_{1min}, \dots, S_{1j}, \dots, S_{1max}\} \text{ avec } j = 0, m_2$$

Le coût global est estimé par Monte Carlo pour chaque point (i,j) du domaine discrétisé selon la procédure suivante:

1. Générer des valeurs aléatoires C and σ à partir des distributions a priori $\pi_C(\cdot)$ and $\pi_\nu(\cdot)$.
2. Générer des instants de défaillance aléatoires et de censure τ à chaque niveau de stress avec les variables aléatoires C et ν obtenues lors de l'étape 1.
3. Estimer C , ν , $V(C)$ et $V(\nu)$ par les relations 5.7 et 5.8
4. Estimer p et $V(p)$
5. Répéter les simulations (de l'étape 1 à 4) k fois (Cette répétition permet de prendre en compte l'effet de taille d'échantillon et d'incertitude sur la fiabilité réelle du nouveau produit caractérisée par les distributions a priori)

6. Estimer la probabilité de défaillance en exploitation en considérant les moyennes p and $V(p)$ sur l’ensemble des répétitions.
7. Evaluer le coût global pour le point (i,j)

Finalement, l’hypersurface du coût global est approximée par un polynôme quadratique. Le plan d’essai optimal est obtenu par une méthode d’optimisation classique en minimisant le coût défini par le polynôme.

| | |
|------------|--|
| C | [400; 28.86] |
| $\pi(C)$ | Normale(400; 28.86) définie par la méthodes des moments |
| ν | [1.3; 1.7] |
| $\pi(\nu)$ | Normale(1.5; 0.1155) définie par la méthodes des moments |

Table 5.3 – Informations a priori sur le modèle SVA

Le tableau 5.3 donne l’information disponible sur les paramètres C et ν . Notons que pour les roulements à billes, le paramètre p est fixé à 1.5. On définit les distributions a priori par deux lois Normales.

Les conditions de sévérisation utilisées pour générer les instants de défaillance et de censure sont définies dans le tableau 5.4.

| | |
|--|---------|
| T temps de censure | 300 h |
| C_k : variable aléatoire tirée de $\pi(C)$ | C_k |
| ν_k : variable aléatoire tirée d $\pi(\nu)$ | ν_k |
| S_2 : Niveau maximal du stress | 175 DaN |
| S_1 : Niveau intermédiaire du stress défini à partir de l’espace a priori discrétisé (p_1, S_1) . | S_1^j |
| n : taille d’échantillon totale | 60 |
| n_1 :taille d’échantillon allouée au niveau S_1 déduite de p_1 définie de l’espace a priori discrétisée (p_1, S_1) . | n_1^i |
| n_2 :taille d’échantillon allouée au niveau S_2 déduite de n_1 par $n_2 = n - n_1$ | n_2^i |

Table 5.4 – Paramètres du plan à optimiser

5.6.2.7 Construction de la surface de réponses

L’espace des solutions discrétisé $(p_1$ et $S_1)$ est défini comme suit:

$$p_1 = \{0.05, 0.15, 0.25, 0.35, 0.45, 0.55, 0.65, 0.75, 0.85, 0.95\}$$

$$S_1 = \{125, 130, 135, 140, 145, 150, 155, 160, 165, 170\}$$

Pour chaque point (i, j) du domaine discrétisé, la probabilité de défaillance est estimée à partir des résultats d’essai simulés selon la méthode présentée précédemment. Cette étape est répétée 500 fois pour prendre en compte les effets de taille d’échantillon et de l’incertitude sur la fiabilité réelle du produit testé caractérisée par les distributions a priori. La probabilité de

défaillance sur la période de garantie est déterminée en considérant la moyenne et la variance tirées des 500 répétitions. Cette approche a été implémentée dans l'environnement MATLAB. Finalement, la fonction coût (essai et exploitation) peut être obtenue. Les paramètres pour évaluer la fonction coût sont donnés dans le tableau 5.5

| | |
|----------------------------|----------------|
| n | 60 équipements |
| Temps de censure | 300 hours |
| Coût d'un équipement testé | 500 |
| Coût fixe d'essai | 10.000 |
| Coût par heure d'essai | 100 |
| Coût par lot | 1000 |
| γ_{essai} | 2.6 |

Table 5.5 – Valeurs des paramètres du Coût d'essai

| | |
|---------------------------------|-------------|
| <i>Pobjectif</i> | 0.01 |
| Population vendue | 1000 unités |
| Période de garantie | 100 heures |
| Coût de rechange | 100 |
| Coût du service après-vente | 10.000 |
| Coût de perte d'image de marque | 100 |
| $\gamma_{exploitation}$ | 1.3 |

Table 5.6 – Valeurs des paramètres de coût d'exploitation

La figure 5.7 représente le coût d'essai dans l'espace des solutions (p_1, S_1) en considérant la fonction coût définie par l'équation et le paramètres donnés le tableau 5.5.

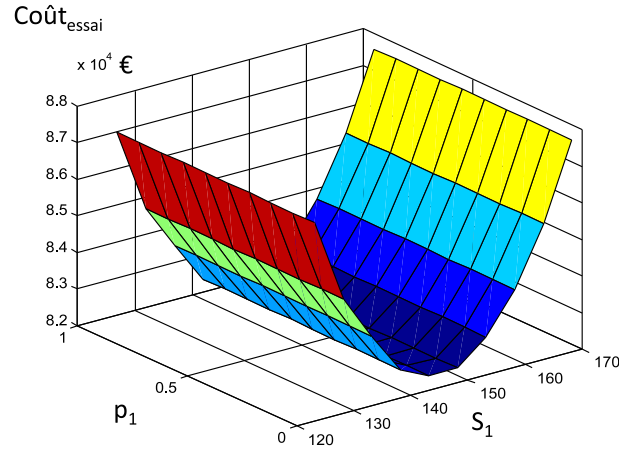


Figure 5.7 – Coût d'essai dans l'espace des solutions $(p_1$ et S_1)

La figure 5.8 représente le coût d'exploitation dans l'espace des solutions (p_1, S_1) en considérant la fonction coût définie par l'équation et le paramètres donnés le tableau 5.6.

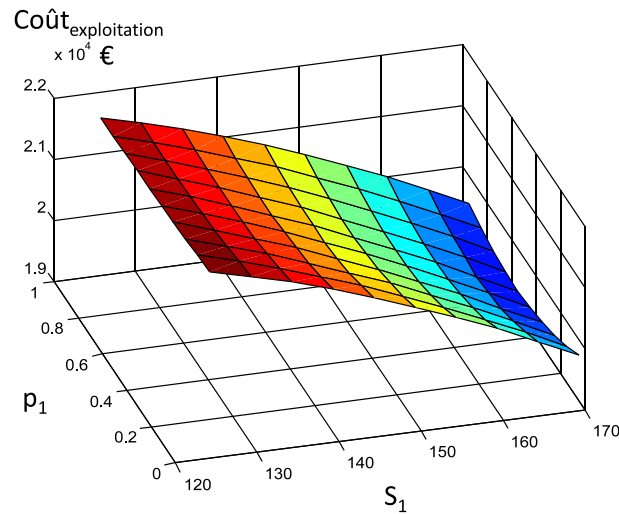


Figure 5.8 – Coût d'exploitation dans l'espace des solutions $(p_1$ et S_1)

La figure 5.9 représente le coût global dans l'espace des solutions (p_1 , S_1).

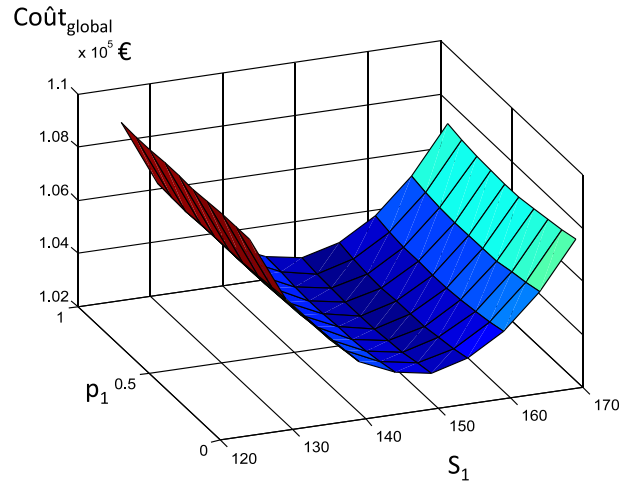


Figure 5.9 – Coût global dans l'espace des solutions (p_1 et S_1)

La figure 5.10 représente les iso-coût global et nous pouvons voir où se situe la zone optimal.

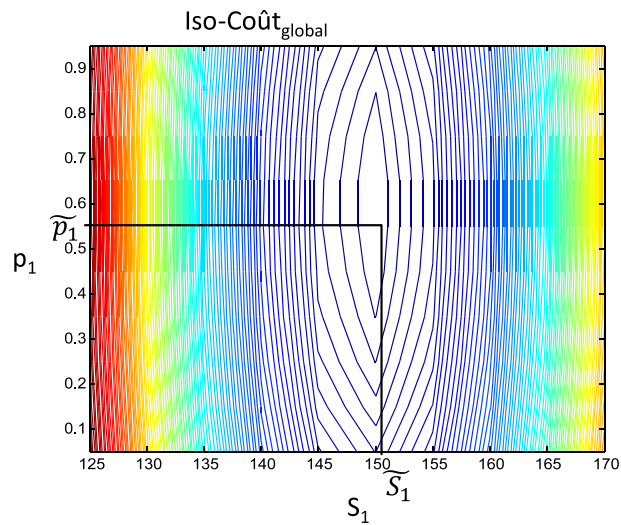


Figure 5.10 – Iso-Coût global dans l'espace des solutions (p_1 et S_1)

Les paramètres du plan de test optimal sont obtenus précisément en approximant la surface du coût global par une surface polynomiale quadratique dont les coefficients sont donnés dans le tableau 5.7.

| | |
|-------|------------|
| a_0 | 119785.123 |
| a_1 | -127.372 |
| a_2 | -27037.498 |
| a_3 | 0.420 |
| a_4 | 25142.345 |
| a_5 | -1.082 |

Table 5.7 – Valeurs des paramètres du polynome

Les paramètres du plan de test optimal sont obtenus finalement par une méthode d’optimisation classique en minimisant le coût décrit par la surface polynomiale. Le minimum est atteint pour $p_1 = 0.541$ et $S_1 = 152.33$ et Coût Global = 102770 . Le tableau 5.8 donne l’ensemble des paramètres du plan optimal.

| | |
|------------------|------------|
| n | 60 unités |
| n_1 | 32 unités |
| n_2 | 28 unités |
| S_1 | 152.33 DaN |
| S_2 | 175 DaN |
| Temps de censure | 300 hours |

Table 5.8 – Paramètres du plan optimisé

5.6.3 Recherche de l’optimum à l’aide des algorithmes génétiques

Dans cette section, nous proposons une procédure d’optimisation utilisant les algorithmes génétiques (GA) pour définir le plan de test optimal. Dans un premier temps, en détaillant notre problématique, nous expliquerons les raisons de ce choix, puis nous proposerons une méthode d’optimisation en conclusion basée sur les algorithmes génétiques. Dans un deuxième temps, nous détaillerons la méthode choisie ainsi que la procédure itérative de recherche de l’optimum.

5.6.3.1 Choix des algorithmes génétiques pour la recherche d’un optimum

La définition d’un plan d’essai accéléré en section 5.5.2 contient des paramètres de décision hétérogènes, discrets ou continus. Les paramètres de décision discrets, c.-à-d. les allocations de composants à tester par niveaux de stress, sont bornés. Le nombre de possibilités est donc fini, toutefois sa combinatoire rend son exploration coûteuse en temps de calcul.

Lorsque le nombre de paramètre de décision devient important, l’utilisation d’une surface de réponse précédemment employé 5.6.1 devient inappropriée en raison de l’explosion de la dimension. Hors, dans la mise en place d’un plan d’essai accéléré, l’ensemble de ses paramètres ω (dont les niveaux de stress S_i , les temps de censure à chaque niveau τ_i , les tailles d’échantillon alouées (n_i) à chaque niveau . . .) peut être considéré comme des paramètres de décision.

Les algorithmes génétiques sont une technique d’optimisation simple et efficace qui peut être appliquée en particulier dans le contexte d’entrées hétérogènes. Le nombre de variables considérées simultanément par les algorithmes génétiques peut aller jusqu’à plusieurs centaines. De plus, la croissance en temps de calcul liée au nombre de variables reste faible en comparaison d’autres méthodes. Il s’agit donc d’une méthode extrêmement flexible qui ne nécessite pas de continuité, de dérivabilité, et de convexité de la fonction objectif à maximiser. Cette flexibilité permet donc de libérer les paramètres de décision voulus d’un plan d’essai en fonction de la problématique sans complexité supplémentaire.

Dans la dernière décennie, ces méthodes ont été introduites dans le cadre d’application statistiques liées à l’ingénierie de la fiabilité. Plusieurs exemples d’applications d’optimisation et d’allocation de la fiabilité d’un système sont régulièrement produits comme [Marseguerra et al., 2006] ou encore [Taboada, 2008]. Dans son éditorial, [Levitin, 2006] fournit une bibliographie des applications des algorithmes génétiques en fiabilité des systèmes. Plus récemment, l’utilité des algorithmes génétiques a été montrée dans les travaux [Hamada et al., 2001] permettant de trouver l’optimum d’un plan d’expérience bayésien pour des modèles de régression. De même, [Lantieri and Guérin, 2012] présente une application à l’optimisation d’un plan d’essai accéléré par pallier de stress.

En conclusion, les algorithmes génétiques fournissent un complément utile à l’expérimentation statistique pour l’optimisation de plan d’essai. Dans la prochaine section, nous détaillerons l’application des algorithmes génétiques à notre problème d’optimisation et les spécificités de celle-ci.

5.6.3.2 Utilisation des algorithmes génétiques

Les algorithmes génétiques appartiennent à la famille des algorithmes évolutionnistes. Leur but est d’obtenir une solution approchée à un problème d’optimisation, lorsqu’il n’existe pas de méthode exacte (ou que la solution est inconnue) pour le résoudre en un temps raisonnable. Ils utilisent la notion de sélection naturelle (sélection, croisements, mutations, etc.) et l’appliquent à une population de solutions potentielles au problème donné. Les algorithmes génétiques ont déjà une histoire relativement ancienne puisque les premiers travaux de John Holland sur les systèmes adaptatifs remontent à 1962 [Holland, 1992].

Un algorithme génétique recherche le ou les extrema d’une fonction définie sur un espace de données. Pour l’utiliser, il est nécessaire de formaliser les cinq éléments suivants :

1. **Un principe de codage de l’élément de la population :** Cette étape associe à chacun des points de l’espace d’état une structure de données. Elle se place généralement après une phase de modélisation mathématique du problème traité. La qualité du codage des données conditionne le succès des algorithmes génétiques associé à la problématique. En effet, c’est sur celui-ci que vont s’appliquer les opérateurs de croisement et de mutation. Les codages binaires ont été très utilisés à l’origine. Les codages réels sont désormais largement utilisés, notamment dans les domaines applicatifs pour l’optimisation de problèmes à variables réelles.
2. **Un mécanisme de génération de la population initiale :** Ce mécanisme doit être capable de produire une population d’individus non homogène qui servira de base pour les

générations futures.

Le choix de la population initiale d’individus conditionne fortement la rapidité de l’algorithme. Si la position de l’optimum dans l’espace d’état est totalement inconnue, il est naturel de générer aléatoirement des individus en faisant des tirages uniformes dans chacun des domaines associés aux composantes de l’espace d’état en veillant à ce que les individus produits respectent les contraintes. En effet, il est essentiel que la population initiale soit répartie sur tout le domaine de recherche afin d’explorer au plus tôt cet espace

3. **Une fonction d’évaluation** : Celle-ci retourne une valeur réelle positive appelée fonction d’évaluation de l’individu. Il s’agit de la fonction à optimiser par l’algorithme.
4. **Des opérateurs permettant de diversifier la population au cours des générations et d’explorer l’espace d’état** : L’opérateur de croisement recompose les gènes d’individus existant dans la population, l’opérateur de mutation a pour but de garantir l’exploration de l’espace d’états et d’éviter la convergence vers des extremums locaux.
5. **Des paramètres de dimensionnement** : taille de la population, contrainte d’optimisation, nombre total de générations ou critère d’arrêt, probabilités d’application des opérateurs de croisement et de mutation.

Le principe général du fonctionnement d’un algorithme génétique est représenté sur la figure 5.11.

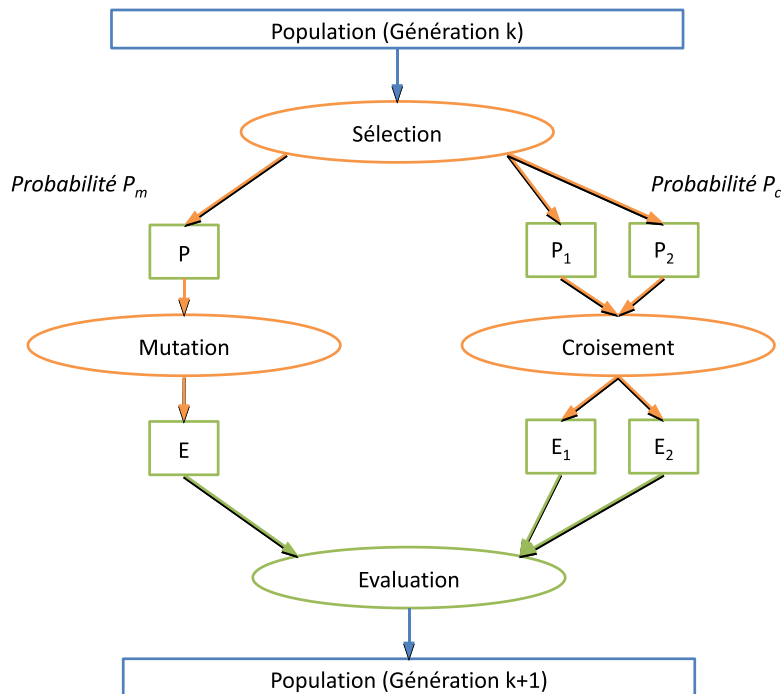


Figure 5.11 – Principe général des algorithmes génétiques

On commence par générer une population d’individus de façon aléatoire. Pour passer d’une

génération k à la génération $k+1$, les trois opérations suivantes sont répétées pour tous les éléments de la population k . Des couples de parents P_1 et P_2 sont sélectionnés proportionnellement à leur aptitude à la sélection naturelle pour la prochaine génération. L'opérateur de croisement leur est appliqué avec une probabilité P_c (généralement autour de 0,6) et génère des couples d'enfants E_1 et E_2 . D'autres éléments P sont sélectionnés en fonction de leur adaptation. L'opérateur de mutation leur est appliqué avec la probabilité P_m (P_m est généralement très inférieur à P_c) et génère des individus mutés E . Le niveau d'adaptation des enfants (E_1, E_2) et des individus mutés E sont ensuite évalués avant insertion dans la nouvelle population. Différents critères d'arrêt de l'algorithme peuvent être choisis :

- Le nombre de générations que l'on souhaite exécuter peut être fixé a priori. C'est ce que l'on est tenté de faire lorsque l'on doit trouver une solution dans un temps limité.
- L'algorithme peut être arrêté lorsque la population n'évolue plus ou plus suffisamment rapidement.

Les algorithmes génétiques nécessitent généralement plus d'une dizaine d'individus pour former une population et à partir de celle-ci, beaucoup de générations pour trouver la solution optimale. Cependant, pour beaucoup d'applications industrielles, l'évaluation d'un individu demande beaucoup de temps de calcul. Dans notre cas, le nombre de simulation de Monte Carlo pour simuler un plan d'essai induit un temps de calcul important. Comme il a été précisé, les choix d'opérateurs et de leurs paramètres associées conditionnent l'efficacité de l'algorithme suivant une problématique précise. Nous allons maintenant exposer les choix effectués sur chacun de ces éléments dans le cas de l'optimisation d'un plan d'essai accéléré, ceci dans le but d'améliorer l'efficacité des algorithmes génétiques, c-à-d. trouver l'optimum en limitant le nombre de générations d'une population.

Codage de la population Un plan d'essai accéléré est défini : Par souci de commodité, la définition d'un individu de la population est restreinte à des variables continues et indépendantes. Ainsi, chaque élément d'une population est défini par une combinaison de facteurs ($/p_i, S_i$) respectant les contraintes du plan d'essai 5.6.

Initialisation Dans le cadre de notre problématique, nous ne connaissons pas la position de l'optimum. Pour constituer un individu, des tirages uniformes sont effectués sur chacun des paramètres du plan. L'individu est sélectionné s'il respecte les contraintes.

Choix de la procédure Dans les algorithmes génétiques classiques, chaque nouvelle génération de la population remplace la précédente. Il peut alors arriver que la meilleure solution de la génération $k + 1$ soit inférieure au sens de son évaluation à la meilleure solution de la génération k . En conséquence, de très bonnes solutions peuvent être perdu. Une solution à ce problème est d'utiliser un algorithmes génétiques dit élitiste. A chaque génération, nous gardons solutions avec les meilleures solutions pour la prochaine génération.

Les performances des algorithmes génétiques dans la recherche d'un optimum global dépendent fortement de l'efficacité des algorithmes de croisement et de mutation. En effet, une proportion trop faible de mutation augmente le risque d'obtenir un optimum local différent de la solution globale. Afin d'améliorer les performances des algorithmes génétiques, notamment l'efficacité des procédures de croisement et de mutation, les algorithmes génétiques adaptatifs sont souvent utilisés dans les études d'optimisation relatives à la fiabilité comme [] .

Au début d'un processus d'algorithme génétique, certains individus peuvent avoir une meilleure évaluation que d'autres. Lorsque le tirage aléatoire est effectué, ces individus sont choisis avec une plus forte probabilités et conduisent à une solution prématurée. Les algorithmes génétiques adaptatifs sont développés pour palier à cet inconvénient. Cependant, il existent de multiples versions d'algorithme génétiques adaptatifs dans la littérature. En suivant les méthodes d'ajustement des taux de mutation et de croisement, les algorithmes génétiques adaptatifs peuvent être classifiés en trois catégories : les algorithmes génétiques adaptatifs donnant la priorité à la mutation, dits "mutation-first", ceux donnant la priorité au croisement, dits "crossover-first" et ceux avec un ordre incertain.

Les approches "mutation-first" mettent l'accent sur l'utilisation d'un taux de mutation plus élevé et un taux de croisement faible au début puis un taux de mutation plus faible et un taux de croisement plus élevé à la fin. Le taux initial élevé de mutation permet d'explorer l'espace des solutions alors que le taux final supérieur de croisement circonscrit la solution optimale.

Les approches "crossover-first" fonctionnent à l'opposé des approches "mutation-first". Elles adoptent un fort taux de croisement au début, puis un fort taux de mutation à la fin. Le taux final élevé de mutation permet d'éviter l'homoplastie (greffe d'un élément de même espece) des individus, pouvant conduire à un problème d'optimum local.

Les algorithmes génétiques adaptatifs sont capables de réduire l'étendue de la population et donc de réduire le temps de simulation. De nouvelles méthodes d'ajustement utilisant la moyenne et l'écart-type d'une population sont employées dans le cas d'algorithmes dit mutation-first et crossover-first. Les travaux de [Ye et al., 2010] montrent que les algorithmes génétiques adaptatifs dit mutation-first sont plus efficaces à notre problématique. Nous les appliquerons donc avec un coefficient de mutation décroissant et un coefficient de croisement croissant proportionnellement au nombre de générations.

Pour résumer, les opérateurs de sélection, de croisement et de mutation choisis sont les suivants :

1. Méthode de sélection :

- (a) Après évaluation de chaque parent par la fonction de coût `??`, les deux meilleurs éléments sont conservés pour la génération suivante (`EliteCount = 2`).
- (b) Ensuite, la sélection des parents pour les opérateurs de croisement se fait uniformément (`SelectionFcn = @selectionstochunif`).

2. Méthode de croisement :

Pour un croisement, deux parents P_1 et P_2 sont choisis suivant une probabilité proportionnelle à leur évaluation. Ils génèrent deux enfants E_1 et E_2 définis par la moyenne de chaque gène des parents. Ainsi, meilleurs les plans P_1 et P_2 sont, plus des enfants similaires sont générés. Lorsqu'il devient difficile de produire une évolution, un optimum local est atteint. La proportion d'enfants issus de croisement, *CrossoverFraction*, augmente au fur et à mesure du nombre de générations *Nbgen*.

$$CrossoverFraction = 0.4 + Nbgen * (0.6/50)$$

3. Méthode de mutation :

Pour une mutation, les deux derniers éléments, i.e les plus mal évalués, sont remplacés par deux autres choisis aléatoirement dans le champ des possibilités. Si les mutations ne conduisent pas à des bons plans, une mauvaise évaluation leur sera attribuée ainsi qu'une faible probabilité d'être choisis par la suite. De plus, la probabilité qu'une mutation intervienne sur chacun des facteurs va décroître inversement au nombre de génération comme le phénomène connu d'"équilibre ponctué". Cet équilibre correspond à un phénomène génétique observable où les mutations s'atténuent au cours du temps avec des sauts dans le processus.

$$\text{mutationRate} = 0.9 * (1 - \text{Nbgen}/50)$$

L'algorithme est implémenté sous MATLAB à partir de la fonction d'algorithme génétique, les options de celle-ci correspondant aux spécificités de la méthode énoncée plus haut.

5.6.4 Exemple numérique pour un plan de qualification

Dans cet exemple numérique, le niveau de stress maximal est fixé au préalable. Il doit être aussi élevé que possible afin de provoquer des défaillances et ainsi diminuer la variance de l'estimation au niveau de stress en condition nominale. Cependant, il ne doit non plus pas être trop élevé pour entraîner des modes de défaillance différents de ceux en condition nominale. Le niveau de stress bas et l'allocation correspondante sont optimisés en minimisant la variance de l'estimation de la moyenne du logarithme de la durée de vie en condition nominale. Le niveau de stress intermédiaire est positionné à égale distance entre le niveau maximum et le niveau bas et l'allocation du nombre de composants à tester est fixée à la moitié de celle au niveau maximum. Dans cet exemple, nous considérons une durée de vie suivant une loi de Weibull avec une seule variable d'accélération sur un module électronique de contrôle d'une pompe. Ce module opère en condition normale à 45°C environ. Pour estimer sa fiabilité en condition nominale, 50 unités doivent être testées à trois niveaux de température. Le niveau haut est positionné à 105°C, correspondant à 5°C en dessous du maximum possible sans modification des modes de défaillance. Les temps de censure sont fixés pour les niveaux de stress bas, intermédiaire et haut respectivement à 1080, 600 et 380 heures.

- La durée de vie X suit une loi de Weibull de paramètre de forme β et de paramètre d'échelle α . DE manière équivalente, le logarithme de la durée de vie $t = \ln(X)$ suit une loi de valeurs extrêmes avec un paramètre d'échelle σ et de location μ tels que $\sigma = \frac{1}{\beta}$ et $\mu = \ln(\alpha)$.
- Le paramètre de forme β ne dépend pas du niveau de stress.
- Le paramètre de location μ est une fonction linéaire de la transformée du niveau de stress S tel que $\mu(S) = \gamma_0 + \gamma_1 \times S$, avec γ_0 et γ_1 les paramètres inconnus à estimer à partir des données d'essai.

Sur les paramètres inconnus de la fonction de fiabilité β , γ_0 et γ_1 , les informations a priori sont :

5.6.4.1 Comparaison des résultats

Yang, ??, détermine le plan d'essai du meilleur compromis qui minimise la variance du maximum de vraisemblance de la moyenne du logarithme de la durée de vie en conditions nominales.

| | | |
|-----------------|------------------|-----------------|
| β_{min} | $\hat{\beta}$ | β_{max} |
| 1,3 | 1,5 | 1,7 |
| γ_{0min} | $\hat{\gamma}_0$ | γ_{0max} |
| -31,65 | -21,65 | -11,65 |
| γ_{1min} | $\hat{\gamma}_1$ | γ_{1max} |
| 7500 | 9980 | 12460 |

Table 5.9 – Informations a priori pour l'exemple

Pour la loi des valeurs extrêmes, la moyenne équivaut au 43^{ème} percentile. La variance du maximum de vraisemblance de la moyenne, noté $\hat{t}_{0,43}$, en conditions nominales est donnée par $Var(\hat{t}_{0,43}) = \frac{\sigma^2}{n} V$ où V est appelé la variance standard.

| | Groupe | 1 | 2 | 3 |
|--|-------------------------|-------|-------|-----|
| Plan du meilleur compromis issu [Yang, 2007] | Nombre d'unités testées | 34 | 5 | 11 |
| | Température (°C) | 74 | 89 | 105 |
| Résultats obtenus avec les mêmes contraintes | Nombre d'unités testées | 34 | 5 | 11 |
| | Température (°C) | 74,54 | 88,98 | 105 |

Table 5.10 – Comparaison des résultats avec ceux de Yang

La table 5.10 compare les résultats des deux approches. Les résultats sont quasi identiques et montre que l'optimisation par les algorithmes génétiques fournit des résultats équivalents.

5.6.4.2 Plan d’essai du meilleur compromis obtenu

Une des principales valeur ajoutée de l’optimisation avec les algorithmes génétiques est de permettre l’ajout d’un grand nombre de variables sans augmenter la complexité et le temps de calcul. Dans cet exemple, le plan de test peut être amélioré en libérant les variables π_2 et S_2 pour la recherche de l’optimum. La table 5.11 donne les résultats obtenus par notre approche en relâchant les contraintes sur π_2 et S_2 .

Les résultats montrent un optimum différent du premier cas. La prise en compte de variables supplémentaires permet donc d’améliorer le coût sur investissement.

| | Coût global | Groupe | 1 | 2 | 3 |
|--|-------------|-------------------------|-------|-------|-----|
| Plan du meilleur compromis issu [Yang, 2007] | 43 364 | Nombre d’unités testées | 34 | 5 | 11 |
| | | Température (°C) | 74 | 89 | 105 |
| Résultats obtenus avec les mêmes contraintes | 43 354 | Nombre d’unités testées | 34 | 5 | 11 |
| | | Température (°C) | 74,54 | 88,98 | 105 |
| Résultats obtenus | 32 265 | Nombre d’unités testées | 9 | 7 | 34 |
| | | Température (°C) | 75,34 | 87,93 | 105 |

Table 5.11 – Résultats avec les algorithmes génétiques et comparaison

5.7 Conclusion

Chacune des méthodes d’optimisation développée présente un intérêt distinct. L’optimisation d’un plan d’essai à l’aide d’une surface de réponse permet une visualisation de l’espace des solutions. Une fois la surface construite, la modification de l’un des paramètres du plan permet de trouver aisément les autres paramètres pour atteindre un optimum. L’optimisation d’un plan d’essai à l’aide d’algorithmes génétiques ne permet pas cette visualisation. Par contre, elle permet d’obtenir un ensemble de solutions optimales lorsque le nombre de variables du plan est important, ceci dans un temps de calcul raisonnable.

Conduite d'un plan d'essais accélérés

6.1 Introduction

Une fois que le plan d'essai est défini, il est mis en oeuvre et les résultats sont collectés afin d'estimer les paramètres du modèle SVA. Ce dernier est ensuite utilisé pour déduire la métrique de fiabilité et vérifier la validité du critère de qualification. Comme nous l'avons précisé dans le chapitre précédent, le critère de qualification est défini par :

$$Prob(m_\alpha \geq m_0) \leq \alpha \quad (6.1)$$

avec m_0 l'objectif de fiabilité à tenir, α le risque fixé et m_α la borne unilatérale.

La méthodologie employée pour définir le plan d'essai optimal et robuste a consisté à définir les paramètres du plan garantissant la prise de décision quant à la qualification quelque soit la fiabilité réelle du produit au risque α donné. Dans le cas où la fiabilité réelle correspond à celle attendue, il est possible de statuer plus rapidement. Aussi, au cours de l'essai il est envisageable de réaliser périodiquement l'estimation des paramètres du modèle SVA et de la métrique de fiabilité pour vérifier le critère de qualification.

Lors de l'analyse, il est possible de se rendre compte que la connaissance a priori sur les paramètres du modèle n'est pas totalement compatible avec les résultats d'essai. Cette problématique a été traitée dans le cadre des thèses d'Emmanuel Usureau et Sorin Voiculescu. Ainsi, il a été défini un coefficient de compatibilité entre l'a priori et la vraisemblance permettant de mesurer le degré de cohérence entre les deux sources d'information (subjective et objective). Ce coefficient peut être employé pour pondérer l'a priori afin de donner plus de poids aux données (vraisemblance).

Ces différentes possibilités peuvent amener l'ingénieur à conduire le plan d'essai selon plusieurs voies. Dans le cadre de la thèse, nous étudierons trois voies :

- l'estimation est effectuée à la fin de la réalisation du plan d'essai en vérifiant la compatibilité entre l'a priori et la vraisemblance pour tester le critère de qualification,
- l'estimation est effectuée périodiquement au cours de l'essai avec vérification de la compatibilité et la qualification est vérifiée pour arrêter le plan plus rapidement,
- l'estimation est effectuée périodiquement au cours de l'essai avec vérification de la compatibilité et le plan d'essai est ajusté (modification des temps de censure, ...) pour tenir compte de la fiabilité réelle du produit et ainsi statuer efficacement sur la qualification (plus rapidement, à coup sûr, ...)

L'ensemble de ces voies est basé sur l'étude de la compatibilité de l'a priori avec les données. Dans les sections suivantes, nous rappelons la méthodologie développée par Sorin Voiculescu ([Voiculescu and Guérin, 2009]).

6.2 Compatibilité de l'a priori avec les données

Parmi les articles sur l'inférence Bayésienne subjective pour estimer le paramètre θ , un grand nombre d'entre eux insiste pour le comportement contrariant du résultat a posteriori quand l'information préalable est mal conditionnée. Ainsi nous considérons une information fautive fournie par les avis experts (ou toute autre source d'information subjective) et une distribution a priori $\pi(\theta)$ dans la limite de la région de confiance de l'approche fréquentielle sur le paramètre θ apportée par des données observées.

Il faut noter que, en cas de données fortement censurées et de petites tailles d'échantillons, quand l'approche Bayésienne informative est très recommandée [Robert, 2001], l'hétérogénéité de données peut avoir les mêmes effets qu'un mauvais a priori avec des données correctes. Un a priori peut être très bien choisi mais l'a posteriori peut impliquer de mauvaises décisions. Par exemple, imaginer un système industriel dont les données de défaillance ont été notées dans le passé. Après quelques amendements le système fonctionne actuellement. Un expert donne le temps probable de défaillance à l'avenir qui peut être décalé, en raison de l'évolution technique. Dans ce cas, certaines données antérieures peuvent être considérées comme une pollution de la connaissance objective. Dans la plupart des cas, les articles sont portés sur un renfort de la robustesse a posteriori. Cependant, une première difficulté est justement de considérer que la connaissance a priori n'est pas contradictoire.

De plus, une connaissance a priori et une vraie connaissance peuvent avoir des rôles symétriques en conflit ; particulièrement dans un contexte industriel, les réponses antérieures de calibrage répondent bien au tremblement des données observées. Curieusement, ce sujet ne semble pas avoir été beaucoup étudié bien qu'il constitue, de notre point de vue, un préliminaire important à l'inférence subjective Bayésienne. Il n'existe pas d'outils simples de statistique pour définir l'accord entre une distribution antérieure et une connaissance objective apportées par des données observées.

Il faut noter dans l'observation stricte du paradigme bayésien que la distribution a priori et la vraisemblance constituent un modèle entier de décision et il semble à première vue non approprié de déterminer les degrés d'anomalie entre eux, d'un point de vue théorique. C'est pourquoi beaucoup d'articles sur l'inférence Bayésienne sont portés sur l'étude de l'accord des données avec le choix du modèle bayésien et la détection de la surprise [Bayarri and Berger, 1998]. À notre connaissance, il y a peu d'articles traitant la détection du conflit entre l'a priori et la vraisemblance. Le principe est de mettre en place des indicateurs habituellement construits pour permettre de modifier l'antérieur ou la probabilité en détectant et en enlevant des données mal positionnées ou des données trop influentes. La question est d'obtenir des a priori robustes permettant un impact a posteriori négligeable quand un conflit est détecté.

[Han, 1997] a développé des méthodes pour construire la distribution antérieure hiérarchique. Récemment, quelques résultats ont été obtenus sur la méthode Bayésienne hiérarchique pour traiter des données de la vie. Mais tous ces résultats obtenus au moyen de méthodes bayésiennes hiérarchiques comportent des intégrations compliquées.

[Dunson and Taylor, 2005] suppose que les données se composent d'un échantillon aléatoire d'une fonction de distribution F_Y , qui est inconnue, et que l'intérêt s'appuie sur les inférences de paramètres données comme un vecteur des quantiles de F_Y . Quand la fonction de probabilité n'est pas entièrement indiquée, une densité postérieure ne peut pas être calculée et l'inférence Bayésienne est difficile. Les auteurs prennent en considération une approche qui se fonde sur une vraisemblance de substitution caractérisée par un vecteur des quantiles.

[Cross et al., 2007] propose une technique Bayésienne pour l'évaluation simultanée de la taille initiale équivalente *flaw size* (EIFS) et les distributions de taux de propagation de la fissure dans les modèles de prévision de la vie basés par mécanique de rupture des données d'inspection. La méthode d'inférence mesure séparément les différentes sources de variabilité pour obtenir une distribution indépendante EIFS sans contraintes pour la connaissance a priori des incertitudes sur d'autres paramètres. La capacité prévoit simultanément et exactement le bruit du taux de propagation de la fissure de distribution et d'EIFS d'une petite taille d'échantillon.

Le théorème de Bayes (equation 4.1) témoigne d'une certaine audace, puisqu'il met sur un

pied d'égalité causes et effets, tous deux pouvant être probabilisés. La différence entre observations et paramètres s'estompe, les manipulations conditionnelles permettant d'interchanger leurs positions respectives.

Quand le paradigme bayésien doit être mis en œuvre pour estimer les paramètres d'un modèle de décision, la modélisation a priori de ces paramètres doit être calibrée très soigneusement. Si la vraisemblance des données observées et la connaissance a priori se révèlent en trop grand désaccord sur la zone de confiance qu'elles induisent pour chacun des paramètres, l'inférence Bayésienne peut conduire à des mauvaises décisions. La détection d'un tel conflit revêt un aspect important - et curieusement sous-estimé - en statistique appliquée à des problèmes concertés.

Dans l'ouvrage [Tribus, 1972], qui constitue une somme des techniques bayésiennes dont la grande émergence se fait à cette époque, il est montré que le χ^2 constitue un exemple de passage à la limite du psi-test (test de plausibilité) bayésien lorsque le nombre de valeurs en présence devient important - ce qui est la condition de travail des statistiques classiques, mais pas nécessairement des bayésiennes. Le raccord entre les deux disciplines, qui sont asymptotiquement convergentes, est ainsi complet.

Le psi-test bayésien (qui est utilisé pour déterminer la plausibilité d'une distribution par rapport à des observations) est asymptotiquement convergent avec le χ^2 des statistiques classiques dans la mesure où le nombre d'observations devient grand. Le choix apparemment arbitraire d'une distance euclidienne dans le χ^2 est ainsi parfaitement justifié a posteriori par le raisonnement bayésien.

[Idée et al., 2001] propose d'employer un test de Fisher entre les mesures antérieures et empiriques de l'incertitude. Aussi, dans le cadre de [Usureau, 2001], on trouve une méthode de vérification des lois a priori modélisées en étudiant la compatibilité avec la vraisemblance ([Usureau et al., 2004], [Guerin et al., 2003], [Guerin et al., 2004]). Ces auteurs ont défini un indicateur évaluant la pertinence de la loi a priori et ont mis en place une méthode permettant une meilleure utilisation de celle-ci par rapport à la distribution de vraisemblance. Le domaine d'existence de cet indicateur est $[0, 1]$. Plus la valeur est proche de 1, plus les deux distributions sont cohérentes. Les auteurs vont plus loin en pondérant l'information a priori avec ce facteur pour diminuer son influence sur l'a posteriori.

L'un des plus récents travaux sur ce sujet a été dirigé dans [Bousquet, 2006]. Dans le cadre de ce rapport, l'auteur propose la validation de l'a priori en employant un critère de cohérence qui mesure la comparabilité entre la modélisation a priori π et l'information apportée par les données. Le critère consiste à calculer le ratio :

$$\text{DAC}^J(\pi | y_n) = \frac{KL(\pi^J(\cdot | y_n) \| \pi)}{KL(\pi^J(\cdot | y_n) \| \pi^J)} \quad (6.2)$$

avec $KL(\pi_1 \| \pi_2) = \int_{\Theta} \pi_1(\theta) \log \frac{\pi_1(\theta)}{\pi_2(\theta)} d\theta$ la divergence Kullback-Leibler entre les distributions π_1 et π_2 et π^J l'a priori non informative. DAC est considéré comme critère de concordance entre les données. Une valeur $\text{DAC}^J(\pi | y_n) \leq 1$ caractérise une concordance entre les données $\cdot | y_n$ et l'a priori π . Une valeur $\text{DAC}^J(\pi | y_n) > 1$ signale un conflit. La méthodologie n'est pas applicable aussi facilement pour des cas plus complexes que les cas idéaux à cause des ajustements nécessaires à faire sur π^J suite à sa impropriété ($\int_{\Theta} \pi(\theta) d\theta \rightarrow \infty$).

Une fois que l'une de ces modélisations a été vérifiée et qu'il est prouvé que l'a priori n'est pas compatible avec l'information apportée par les données, trois solutions peuvent être envisagées :

- revenir à l'utilisation des méthodes classiques, ce qui implique soit la perte d'intérêt pour les méthodes Bayésiennes, soit la disparition d'une partie d'information qui se traduit dans des pertes économiques,
- recommencer la modélisation de la connaissance a priori, ce qui peut se révéler long et coûteux en terme d'analyse de données ou d'interrogation des experts,
- pondérer la connaissance a priori pour la rendre compatible avec les résultats réellement obtenus (vraisemblance) et donc ne retenir qu'une partie des informations a priori. En effet, même mal modélisée, la connaissance a priori est en partie basée sur des données correctes qui peuvent être exploitées dans le cadre de l'inférence Bayésienne.

6.2.1 Le principe de calcul d'un facteur de compatibilité normalisé

Dans le cadre Bayésien, la probabilité retrouve sa sémantique naturelle: évaluer l'incertitude allant du connu (les données statistiques) vers l'inconnu (les hypothèses d'intérêt). En combinant les données avec une distribution initiale qui synthétise les informations autres que les données, la règle de Bayes conduit à une distribution finale qui permet de calculer les probabilités des hypothèses.

L'inférence Bayésienne classique dans le cadre de l'analyse de la fiabilité expérimentale, par la formulation mathématique donnée par l'équation 4.1, en faisant le produit direct entre la fonction a priori et la vraisemblance, donne le même poids aux deux distributions : la vraisemblance et la connaissance a priori.

Pour les cas où l'une des deux serait trop informative, on observe que l'autre ne compte plus.

Dans le cadre de la démarche présentée dans [Usureau, 2001], il est montré une approche qui permet de pondérer l'information a priori en fonction du degré de compatibilité entre l'a priori et la vraisemblance. Nous proposons d'étudier l'application de cette approche à notre problématique.

Cette méthodologie permet de déterminer un facteur K de pondération normalisé compris entre 0 et 1 qui permet :

- de déterminer le degré de compatibilité de la modélisation de la connaissance a priori avec les données réelles,
- de pondérer la distribution a priori afin de la rendre compatible avec les données observées et pour ne pas biaiser les résultats de l'analyse.

Un tel facteur de pondération doit pouvoir être utilisé dans le cas général, exploitable et efficace dans tous les cas extrêmes possibles.

La méthodologie compare l'information a priori qu'on a pour chaque paramètre θ_i avec l'information contenue par la vraisemblance définie par les fonctions marginales (voir section 2.3.2).

Ce facteur doit mesurer la similitude entre la distribution a priori initiale modélisée à partir des avis d'experts, des données et/ou simulations sur des produits similaires et la fonction de vraisemblance représentant, elle, les données réelles observées.

Cette similitude est proportionnelle au produit de convolution (voir Figure 6.1) entre l'a priori $g_0(\theta)$ et la vraisemblance $f(\theta)$. Il est ainsi possible de définir le réel $C(\delta)$ par le produit de convolution $g_0(\theta) \otimes f(\theta)$, soit :

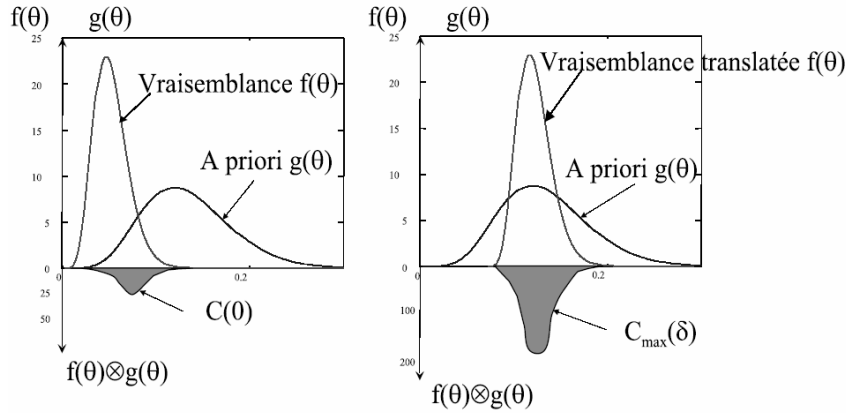


Figure 6.1 – Principe de détermination du coefficient de pondération normalisé

$$C(\delta) = \int_{D(\theta)} g_0(\theta) f(\theta - \delta) d\theta \quad (6.3)$$

Certes, on peut procéder d'une autre manière en écrivant $C(\delta) = \int_{D(\theta)} g_0(\theta - \delta) f(\theta) d\theta$, cela ne change en rien le calcul.

La valeur de $C(\delta)$ évolue bien avec la similitude entre la distribution a priori $g_0(\theta)$ et la distribution de vraisemblance $f(p)$ puisque plus les deux fonctions se recoupent et plus $C(\delta)$ est élevée. Dans le cas contraire, plus les deux fonctions ne sont pas similaires et plus $C(\delta)$ est faible.

Il reste maintenant à trouver une solution pour réaliser la pondération de la distribution a priori. Pour ce faire, il suffit de définir un valeur cohérente avec $C(\delta)$ qui évolue uniquement entre 0 et 1 . L'expression normalisée de $C(\delta)$ suffit à répondre à ce problème, notons la K :

$$K = \frac{C(0)}{C_{\max}(\delta)} \quad (6.4)$$

avec

- $C(0)$ obtenue pour $\delta = 0$
- $C_{\max}(\delta)$ représentant la valeur maximum de $C(\delta)$, c'est-à-dire pour simplifier le maximum de l'aire qui se trouve sous la courbe produit des deux fonctions $g_0(\theta)$ et $f(\theta)$ quand on fait glisser les deux courbes l'une sur l'autre (voir figure 6.1).

Finalement, K n'est autre qu'un facteur de compatibilité normalisé entre la distribution a priori et la distribution de vraisemblance.

6.2.2 L'utilisation du facteur de compatibilité comme facteur de pondération de l'a priori

Nous disposons maintenant d'un facteur K de pondération qui permet de statuer sur la bonne cohérence entre la modélisation a priori et les essais réels symbolisés par la vraisemblance. Bien entendu, le facteur K est déjà utile dans le sens où il va permettre de statuer sur la pertinence de l'a priori.

Il est possible d'aller plus loin dans la démarche. En effet, quel que soit l'état de la distribution a priori, le facteur K peut être utilisé pour pondérer cette connaissance afin d'améliorer la compatibilité avec la vraisemblance.

Plus K sera proche de 1 et plus la connaissance a priori originale sera intégrée, permettant ainsi d'exploiter au mieux l'un des intérêts des méthodes bayésiennes, à savoir de réaliser des gains considérables sur les plans d'essais de Sécurité/Fiabilité.

Plus K sera proche de 0, moins la connaissance a priori initiale sera intégrée dans l'inférence Bayésienne, évitant ainsi de prendre des décisions pouvant être lourdes de conséquences sur le plan de la Sécurité ou sur le plan économique, ceci sur la base de résultats faux.

Le but ainsi serait de pondérer une distribution de telle façon que la distribution pondérée soit $1/K$ fois moins informative.

L'information de Fisher est une notion qui quantifie l'information relative à un paramètre contenu dans une distribution. Sa définition est fortement liée à la variance 2.26. On interprète la variance comme l'espérance des carrés des écarts à l'espérance. Elle permet de caractériser la dispersion des valeurs par rapport à la moyenne. Ainsi, une distribution avec une même espérance et une variance plus grande apparaîtra comme plus étalée.

Le résultat est qu'on ne modifie pas le mode de la distribution a priori, par contre la variance devient plus importante affaiblissant cette distribution. Le facteur de pondération K évoluant avec le niveau de pondération entre a priori et vraisemblance, la modélisation a priori initiale sera d'autant plus intégrée dans l'inférence Bayésienne qu'elle est compatible avec la vraisemblance. Il en sera de même en cas d'incompatibilité entre l'a priori et la vraisemblance car le facteur K sera proche de 0. Le poids de la vraisemblance dans l'inférence Bayésienne sera donc prépondérant. Aussi, la modélisation a priori incompatible au départ sera très peu exploitée.

Ainsi, le premier moment de la nouvelle distribution serait identique avec l'ancien mais le deuxième moment serait $1/K$ fois plus grand.

Pour le calcul pratique nous distinguons deux cas possibles :

- Le cas où l'on peut trouver la solution en résolvant directement les égalités entre les équations décrivant les moments d'ordre 1 et 2.

Par exemple, considérons le cas d'une distribution a priori $f(x)$ de type gamma de paramètres α et λ , de densité de probabilité $f(x) = x^{\alpha-1} \frac{e^{-\frac{x}{\lambda}}}{\Gamma(\alpha)\lambda^\alpha}$, avec l'espérance $\alpha\lambda$ et la variance $\alpha\lambda^2$. La pondération de la distribution a priori par le facteur K donne une nouvelle distribution de paramètres tels que :

$$\begin{cases} \alpha\lambda = \alpha'\lambda' \\ \alpha\lambda^2 = K\alpha'\lambda'^2 \end{cases} \quad (6.5)$$

Qui donnent $\alpha' = K\alpha$ et $\lambda' = \frac{\lambda}{K}$.

- Le cas où, pour résoudre les égalités précédentes, il faut appliquer une méthode numérique d'optimisation.

Par exemple, considérons le cas d'une distribution a priori $f(x)$ de type Weibull de paramètres β et η , de densité de probabilité $f(x) = \frac{\beta}{\eta} \left(\frac{x}{\eta}\right)^{\beta-1} e^{-\left(\frac{x}{\eta}\right)^\beta}$, avec l'espérance $\eta\Gamma\left(1 + \frac{1}{\beta}\right)$ et la variance $\eta^2\Gamma\left(1 + \frac{1}{\beta}\right) - \left(\eta\Gamma\left(1 + \frac{1}{\beta}\right)\right)^2$. La pondération de la distribution a priori par le facteur K donne une nouvelle distribution de paramètres vérifiant :

$$\begin{cases} \eta\Gamma\left(1 + \frac{1}{\beta}\right) = \eta'\Gamma\left(1 + \frac{1}{\beta'}\right) \\ \eta^2\Gamma\left(1 + \frac{1}{\beta}\right) - \left(\eta\Gamma\left(1 + \frac{1}{\beta}\right)\right)^2 = K\eta'^2\Gamma\left(1 + \frac{1}{\beta'}\right) - \left(\eta'\Gamma\left(1 + \frac{1}{\beta'}\right)\right)^2 \end{cases} \quad (6.6)$$

Les paramètres β' et η' sont obtenus en résolvant numériquement ces équations.

A titre d'illustration, les figures suivantes présentent des exemples d'inférence Bayésienne utilisant une distribution a priori pondérée dans la construction de celle correspondant à l'a posteriori en considérant les systèmes à mortalité exponentielle (à gauche) et binômiale (à droite).

Distributions a priori très pessimiste ou optimiste => K très faible

Dans les figures ci-dessous, que la distribution a priori soit très optimiste (figure 6.2) ou bien très pessimiste (figure 6.3), le facteur de pondération K est très faible.

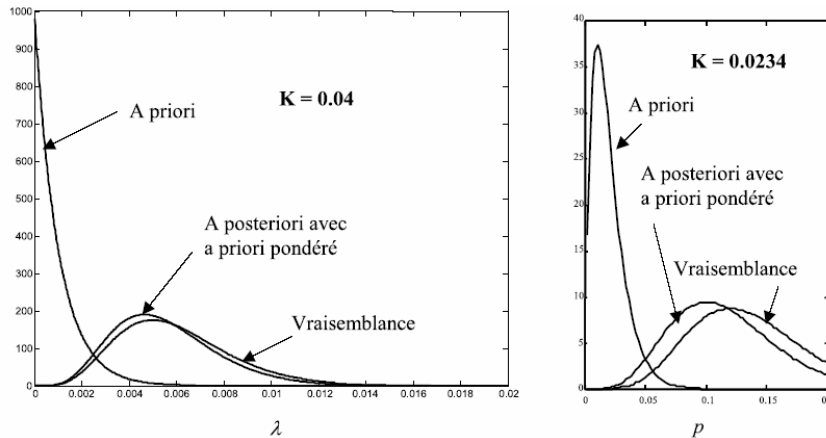


Figure 6.2 – Distribution a priori décalée et trop optimiste

La pondération, par ce facteur, de la distribution a priori permet de donner un maximum de poids à la vraisemblance.

La distribution a posteriori est donc modélisée principalement en exploitant les résultats de la vraisemblance puis, à une moindre échelle, par les données cohérentes de l'a priori. Les résultats sont donc plus proches de la réalité et l'inférence Bayésienne a mieux été optimisée.

Distribution a priori non cohérente => K minimum

Dans le cas de la figure 6.4, la distribution a priori n'est absolument pas cohérente et beaucoup trop optimiste. La pondération par K , qui cette fois-ci est nul, permet de ne pas prendre en compte l'erreur liée à l'a priori. La distribution a posteriori est identique à la distribution de vraisemblance, ce cas correspond au cas des essais classiques.

Distribution a priori cohérente => K maximum

Dans le cas de la figure 6.5, la pondération entre a priori et la vraisemblance est maximale. La pondération par le facteur K égal à 1 n'a aucun effet sur l'a posteriori.

Distribution a priori en partie cohérente => 0 < K < 1

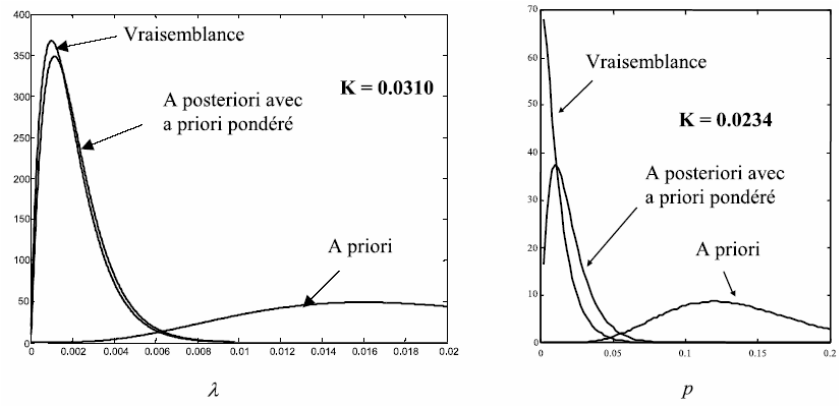


Figure 6.3 – Distribution a priori décalée et trop pessimiste

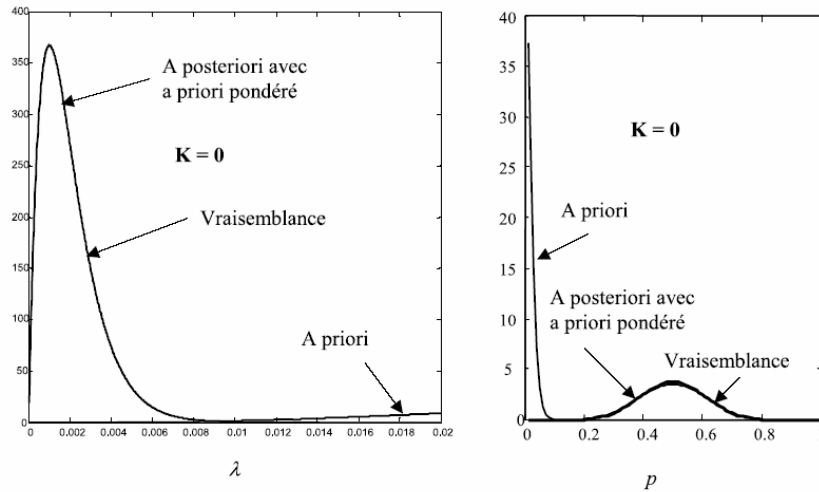


Figure 6.4 – Distribution a priori incohérente

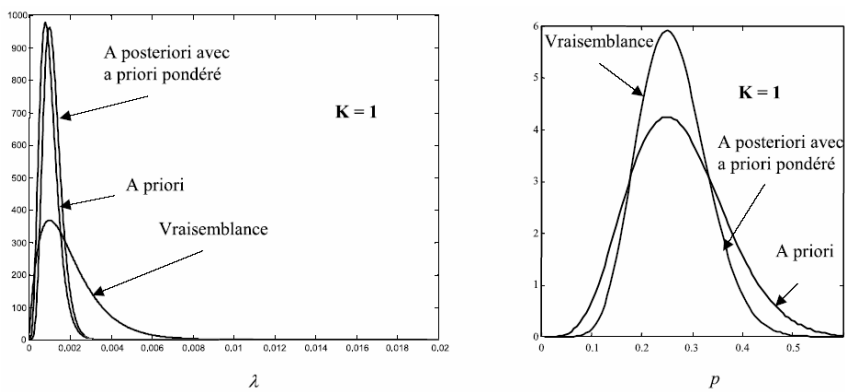


Figure 6.5 – Distribution a priori cohérente

Le cinquième cas (figure 6.6) est une illustration de toutes les éventualités comprises lorsque K n'est ni minimum, ni maximum. Dans ce cas, seule la partie cohérente de l'a priori est prise en compte dans l'inférence Bayésienne, laissant plus ou moins de poids à la vraisemblance en fonction de la similitude des deux fonctions. Ceci permet d'inhiber les erreurs liées à la première modélisation a priori et de fournir des résultats fiables à l'analyste et au décideur.

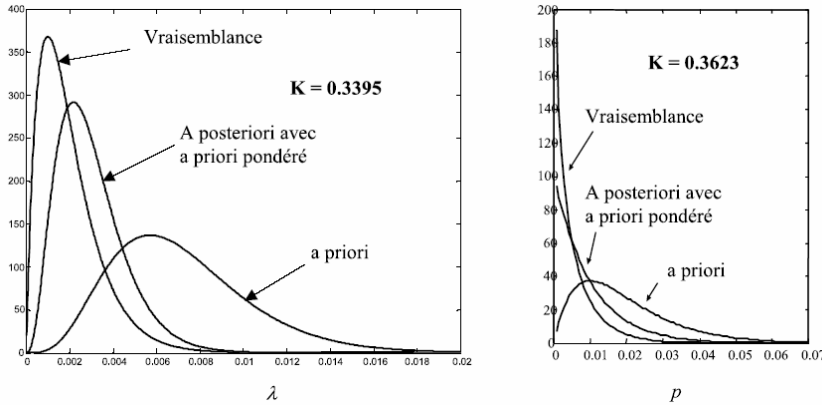


Figure 6.6 – Distribution a priori partiellement cohérente

6.2.3 Application au modèle SVA

La recherche de la compatibilité entre la connaissance a priori et la vraisemblance consiste à déterminer les facteurs K pour chacun des paramètres du modèle SVA considéré (paramétrique ou semi-paramétrique). Pour ce faire, nous déterminons les fonctions marginales ($g_\theta(\theta)$) associées aux paramètres θ ($f_{\theta_i}(\theta_i)$ pour le paramètre θ_i) du modèle par :

$$f_{\theta_i}(\theta_i) = \frac{\int_{j=1 \dots q, j \neq i} \dots \int L(\theta_1, \dots, \theta_q) \cdot \prod_{j=1 \dots q, j \neq i} d\theta_j}{\int_{j=1 \dots q} \dots \int L(\theta_1, \dots, \theta_q) \cdot \prod_{j=1 \dots q} d\theta_j} \quad (6.7)$$

Ensuite, nous déterminons le facteur K_{θ_i} selon la méthode présentée dans le paragraphe 6.2.1 en considérant les distributions marginales $f_{\theta_i}(\theta_i)$ et a priori $\Pi_{\theta_i}(\theta_i)$. Une fois les facteurs K_{θ_i} estimés, nous déterminons les distributions a priori pondérées d'après la méthode détaillée dans le paragraphe précédent. Cette étape n'est pas systématique, car l'incompatibilité peut s'expliquer par d'autres causes qu'un problème d'a priori. Ainsi, nous obtenons la nouvelle formulation de l'inférence bayésienne :

$$\pi_{apo}(\theta/t_i) = \frac{L(t_1, \dots, t_n/\theta) \prod_{i=1}^q \pi(\theta_i, K_{\theta_i})}{\int_{\Omega} L(t_1, \dots, t_n/\theta) \prod_{i=1}^q \pi(\theta_i, K_{\theta_i}) d\theta_i} \quad (6.8)$$

avec comme notations $L(t_1, \dots, t_n/\theta)$ la vraisemblance (définie par les équations 4.2), $\pi(\theta_i, K_{\theta_i})$ la distribution a priori pondérée pour le paramètre θ_i et $\pi_{apo}(\theta/t_i)$ la fonction a posteriori.

Pour illustrer cette technique de pondération, considérons l'exemple présenté dans le paragraphe 4.4.3 en étudiant uniquement le paramètre v . La figure 6.7 présente les fonctions de densité de l'a priori (courbe 1) et de la vraisemblance (courbe 2) et l'a posteriori (courbe 3). On observe que l'a posteriori est décalé par rapport à la valeur théorique et quasiment superposé à l'a priori. Cela s'explique par le choix d'un a priori incohérent et très informatif.

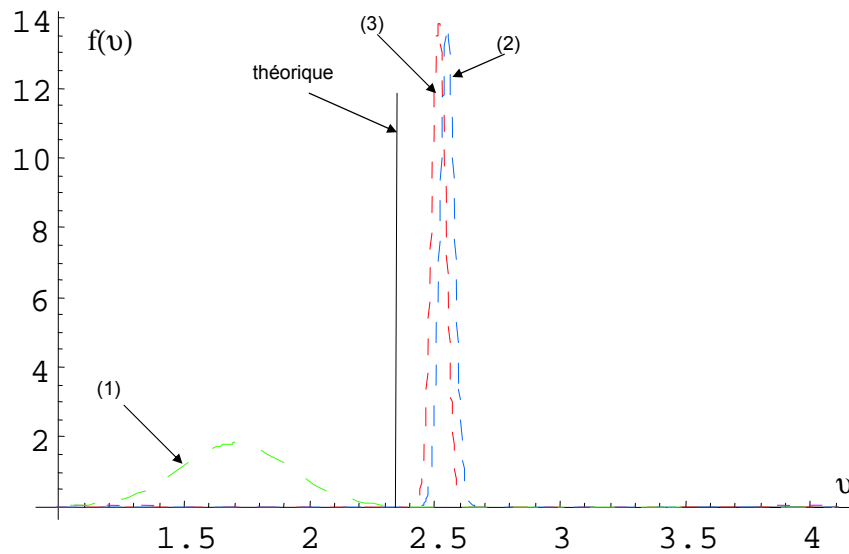


Figure 6.7 – Estimation classique pour ν

L'application du critère DAC ([Bousquet, 2006]) donne une valeur de 1.4 ce qui invalide le choix d'a priori pour le jeu de données. Le calcul de facteur de pondération K donne une valeur de 0.135, valeur qui est suffisamment petite pour invalider aussi le choix d'a priori. En plus, en allant plus loin, la pondération de l'a priori avec le facteur $K = 0.135$ donne la courbe (4) de la figure 6.8. L'application de démarche Bayésienne sur l'a priori ainsi pondérée donne l'estimation a posteriori (courbe 5), estimation moins informative mais proche de la valeur théorique.

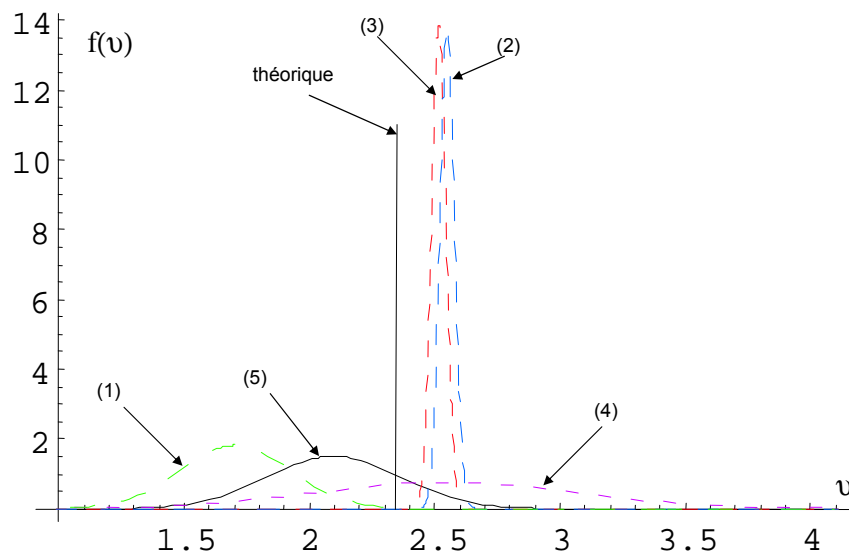


Figure 6.8 – Pondération pour ν

6.3 Analyse finale ou périodique des résultats d'essai

Ces deux cas de figure reprennent en grande partie l'analyse présentée précédemment. La différence concerne la vraisemblance qui intégrera l'ensemble des résultats d'essai ou évoluera au cours de l'essai afin d'intégrer les nouveaux instants de défaillance et de censure. Ainsi, à une date d'analyse, on relève les différents résultats. Il est imaginable, en fonction des contraintes industrielles, que certains niveaux d'essai ne soient pas encore réalisées (car, par exemple, il y a une seule machine). On recherche la compatibilité entre la connaissance a priori et la vraisemblance en déterminant les facteurs K pour chacun des paramètres du modèle SVA considéré selon la même méthodologie.

$$f_{\theta_i}(\theta_i) = \frac{\int_{j=1\dots q, j \neq i} \dots \int L(\theta_1, \dots, \theta_q) \cdot \prod_{j=1\dots q, j \neq i} d\theta_j}{\int_{j=1\dots q} \dots \int L(\theta_1, \dots, \theta_q) \cdot \prod_{j=1\dots q} d\theta_j} \quad (6.9)$$

Ensuite, nous déterminons le facteur K_{θ_i} selon la méthode présentée dans le paragraphe 6.2.1 en considérant les distributions marginales $f_{\theta_i}(\theta_i)$ et a priori $\Pi_{\theta_i}(\theta_i)$. Une fois les facteurs K_{θ_i} estimés, nous déterminons les distributions a priori pondérées d'après la méthode détaillée dans le paragraphe précédent. Cette étape n'est pas systématique, car l'incompatibilité peut s'expliquer par d'autres causes qu'un problème d'a priori. Ainsi, nous obtenons la nouvelle formulation de l'inférence bayésienne :

$$\pi_{apo}(\theta/t_i) = \frac{L(t_1, \dots, t_n/\theta) \prod_{i=1}^q \pi(\theta_i, K_{\theta_i})}{\int_{\Omega} L(t_1, \dots, t_n/\theta) \prod_{i=1}^q \pi(\theta_i, K_{\theta_i}) d\theta_i} \quad (6.10)$$

avec comme notations $L(t_1, \dots, t_n/\theta)$ la vraisemblance (définie par les équations 4.2) $\pi(\theta_i, K_{\theta_i})$ la distribution a priori pondérée pour le paramètre θ_i et $\pi_{apo}(\theta/t_i)$ la fonction a posteriori.

Par la méthode MAP, on détermine les paramètres du modèle SVA $\hat{\theta} = \{\hat{\gamma}, \hat{\sigma}\}$. Ensuite on détermine la matrice de Fisher $I^{MAP}(\hat{\gamma}, \hat{\sigma})$ pour déduire la matrice de Variance-Covariance $\Sigma(\hat{\gamma}, \hat{\sigma})$. Finalement, nous déterminons la métrique de fiabilité et on vérifie le critère de qualification. Dans le cas où le critère de qualification n'est pas vérifié, on poursuit le plan d'essai initial.

6.4 Analyse périodique des résultats et ajustement du plan d'essai

Cette dernière approche de suivi de plan d'essai reprend toute l'analyse présentée dans la section précédente mais dans le cas où le critère de qualification n'est pas vérifié, on propose de rechercher de nouveaux paramètres optimaux de plan. Évidemment, ce nouveau plan $P(\omega^*)$ reprend en partie la configuration initiale. En effet, les contraintes industrielles ne permettent pas de repartir sur un nouveau plan mais par contre il est possible d'agir que sur certains paramètres ω^* du plan initial $P(\omega)$ (comme par exemple les temps de censure, les niveaux de stress lorsque les essais n'ont pas commencé, ...).

Ainsi, le problème d'optimisation se formule de la manière suivante :

Rechercher les paramètres ω^* du plan ajusté $P(\omega^*)$ tels que

$$\left| \begin{array}{l} \text{Minimiser } C(\omega^*, \vartheta, m(\gamma, \sigma), m_0, \alpha) \\ \text{sous contraintes : } g_i(\omega_i) \geq 0, i = 1, \dots, r \\ \gamma_l^{min} \leq \gamma_l \leq \gamma_l^{max}, l = 1, \dots, m \\ \sigma^{min} \leq \sigma \leq \sigma^{max} \end{array} \right. \quad (6.11)$$

L'implémentation du problème d'optimisation s'effectue selon l'approche par algorithme génétique.

6.5 Exemple numérique de l'analyse périodique des résultats d'essai

En reprenant l'exemple 5.6.4, nous définissons une stratégie de suivi de plan permettant d'examiner la pertinence de poursuivre le plan initial ou de proposer une modification de plan en prenant en compte les premiers résultats. Pour ce faire, nous suivons la méthodologie suivante :

1. Estimation de la métrique de fiabilité avec les premiers résultats d'essai et estimation des facteurs de compatibilité entre les fonctions marginales des paramètres inconnus et les distributions a priori associées.
2. Selon le niveau des facteurs de compatibilité, nous proposons :
 - K_θ proche de 1, on teste le critère de qualification. Dans le cas où le test est positif alors on arrête l'essai sinon on poursuit le plan selon la définition initiale.
 - K_θ proche de 0, on détermine de nouveaux paramètres de plan optimisant le coût en prenant en compte qu'une partie du plan initial a été réalisée. Dans ce cas seuls certains paramètres peuvent faire l'objet d'une optimisation.
 - K_θ entre 0 et 1, on choisit la meilleure stratégie en réalisant le test de qualification et la recherche de nouveaux paramètres de plan.
3. Poursuite de l'essai selon la stratégie choisie jusqu'à la prochaine étape d'analyse où les points 1 et 2 sont de nouveau testés.

6.5.1 Stratégie de conduite de l'essai

Pour la conduite du plan d'essai présenté à l'exemple 5.6.4, plusieurs hypothèses sont prises pour définir la stratégie de mise en place des essais.

Hypothèses :

- Un seul banc d'essai, ce qui implique que les essais sur les différents niveaux de stress sont effectués à la suite.
- La conduite des essais commence par le niveau de stress le plus élevé, puis va décroissant vers les niveaux de stress les plus faibles. Cela permet d'observer le maximum de défaillance au départ pour un temps minimum afin de vérifier la compatibilité. Ici, il peut y avoir plusieurs stratégies, mais celle-ci semble la plus adaptée.
- Les résultats intermédiaires d'essai sont exploités une fois le temps de censure terminé pour un niveau de stress donné sur un banc d'essai.

Le déroulement des essais sont représentés par la séquence schématisée sur la figure 6.9.

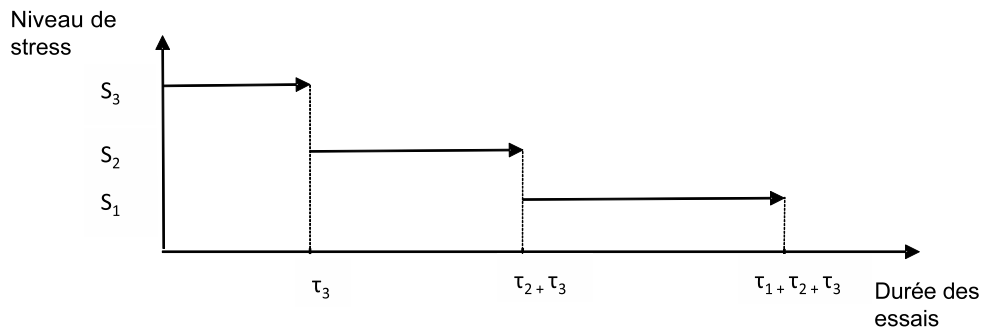


Figure 6.9 – Séquence de réalisation des essais aux différents niveaux au cours du temps

Dans les paragraphes suivant, nous présentons différentes situations possibles.

6.5.2 Cas 1 : Acceptation du test à la première étape

Cette situation concerne le cas d'acceptation dès le premier palier réalisé, il est possible de statuer sur la qualification du produit.

Les valeurs initiales choisies pour la simulation du plan sont :

$$\gamma_0 = -21,65,$$

$$\gamma_1 = 10000,$$

$$\beta = 1,5.$$

Il s'agit des valeurs moyennes de chaque paramètres issues des informations a priori.

Les résultats sont donnés sur la figure 6.10 suivante.

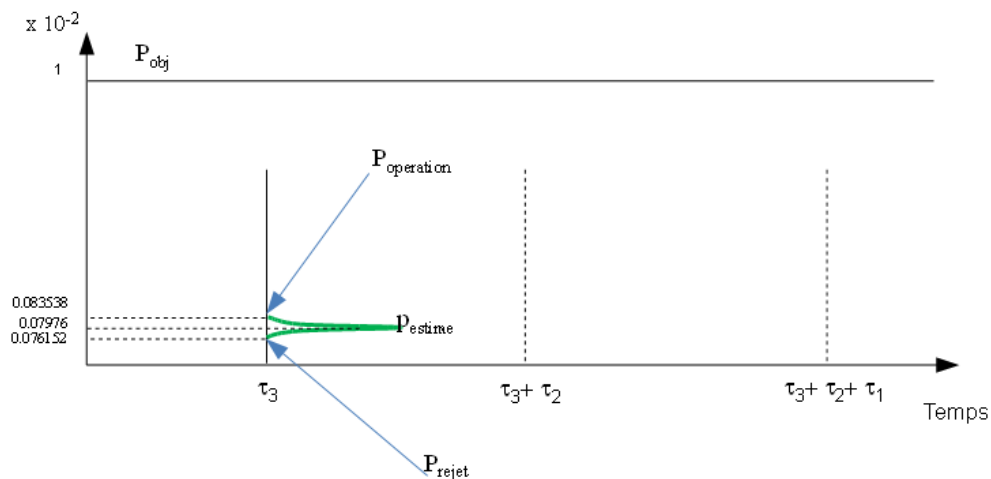


Figure 6.10 – Cas 1 : Acceptation du test à la première étape

La distribution représentée sur la figure caractérise la probabilité de défaillance estimée à partir des données simulées et des informations a priori. La droite horizontale à 0.01 marque la valeur objectif et nous pouvons constater que l'intervalle de confiance bilatéral symétrique (au niveau de risque $\alpha = 0.05$) tiré de la distribution de Pf est largement en dessous du seuil. La borne inférieure (P_{rejet}) permet dans certaine situation de rejeter la qualification. Il est possible de construire un intervalle de confiance dissymétrique pour prendre en compte des valeurs de risque client et fournisseur différentes. Les essais sont réalisés au niveau de stress le plus élevé pour provoquer rapidement des défaillances. Dans ce cas, nous obtenons des informations précoces de nos paramètres permettant une acceptation de la qualification du produit de manière anticiper. L'anticipation de la décision de qualification permet ici de préserver une partie des coûts d'essai.

6.5.3 Cas 2 : Acceptation du test à la deuxième étape

Cette deuxième situation concerne le cas d'acceptation au deuxième palier réalisé. Là encore, il est possible de statuer sur la qualification du produit avant la fin des essais.

Les valeurs initiales choisies pour la simulation du plan sont :

$$\gamma_0 = -27,$$

$$\gamma_1 = 11500,$$

$$\beta = 1,6.$$

Les résultats sont donnés sur la figure 6.11 suivante.

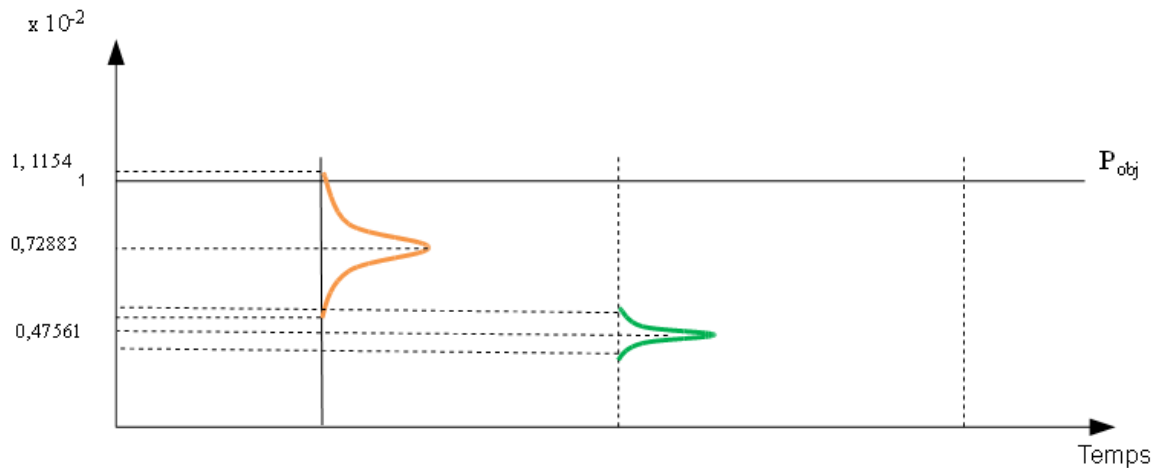


Figure 6.11 – Cas 2 : Acceptation du test à la deuxième étape

Dans ce cas, nous obtenons des informations précoces de nos paramètres permettant une acceptation de la qualification du produit de manière anticiper à la deuxième étape. L'anticipation de la décision de qualification permet là encore de préserver une partie des coûts d'essai.

6.5.4 Cas 3 : Refus du test à la troisième étape

La troisième situation concerne un cas de refus à la fin de la réalisation des essais. Dans ce cas, la décision de refus de la qualification ne peut être prise avant la fin des essais.

Les valeurs initiales choisies pour la simulation du plan sont :

$$\gamma_0 = -27,$$

$$\gamma_1 = 11000,$$

$$\beta = 1,6.$$

Les résultats sont donnés sur la figure 6.12 suivante.

Dans ce cas, l'analyse intermédiaire ne permet pas d'anticiper le résultat quant à la décision de qualification. Une optimisation du plan d'essai à la deuxième étape permettrait peut-être

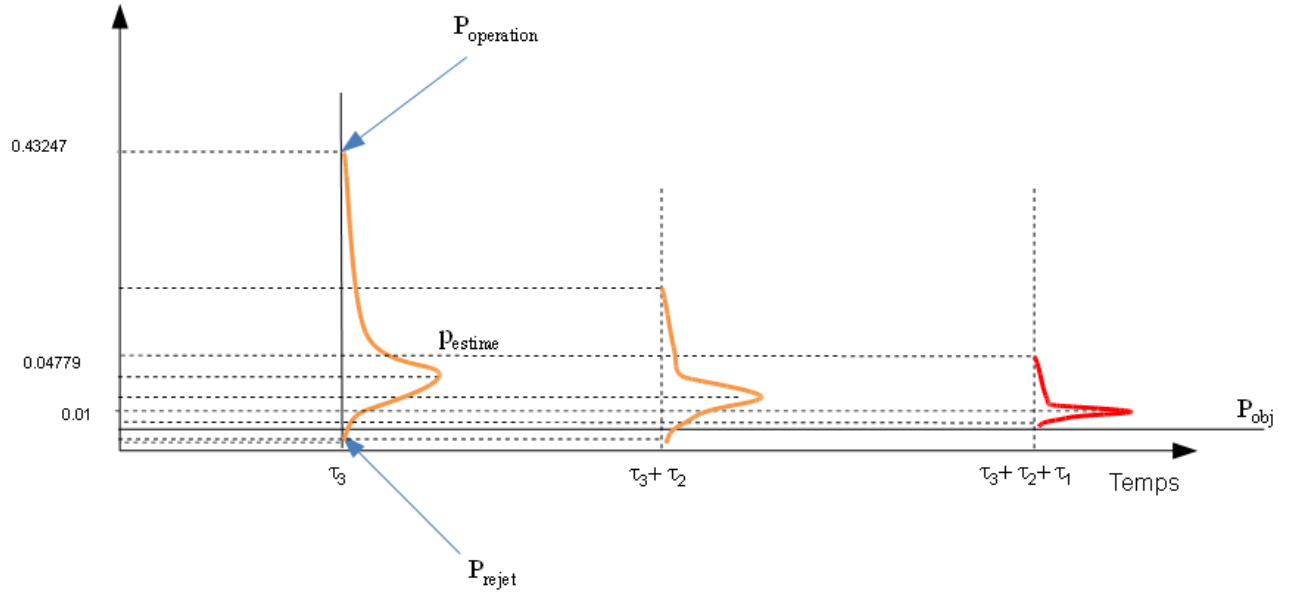


Figure 6.12 – Cas 3 : Refus du test à la 3ème étape

d'augmenter le nombre d'éléments testés au niveau de stress intermédiaire afin de provoquer des défaillances plus rapidement et améliorer la précision de l'estimation à la deuxième étape.

6.5.5 Cas 4 : Acceptation du test à la troisième étape

Les valeurs initiales choisies pour la simulation du plan sont :

$$\gamma_0 = -26, 2,$$

$$\gamma_1 = 11000,$$

$$\beta = 1, 6.$$

Les résultats sont donnés sur la figure 6.13 suivante.

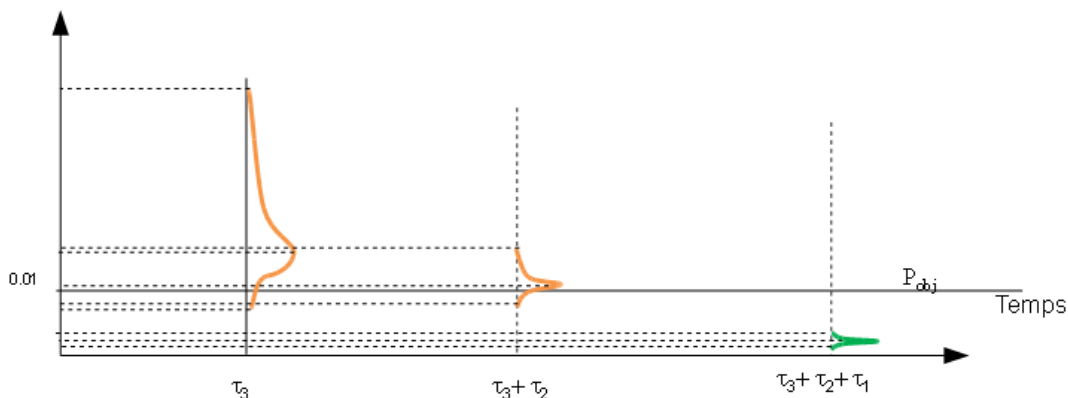


Figure 6.13 – Cas 4 : Acceptation du test à la 3^{ème} étape

Ce quatrième cas, comme le précédent, montre l'utilité de la nécessité d'une optimisation intermédiaire du plan d'essai.

6.6 Conclusion

Dans ce chapitre, nous nous sommes intéressés au suivi lors de la réalisation des essais. En effet, la méthodologie présentée au chapitre précédent peut être réemployée pour anticiper la décision de qualification et améliorer le coût des essais.

Pour permettre ce suivi, il est d'abord d'utiliser un facteur de compatibilité avec les données a priori. Ce facteur permet d'estimer la fiabilité à chacune des étapes de la réalisation du plan en cohérence entre les données d'essai et les informations a priori.

Cette utilisation est illustrée sur un exemple avec différents cas de données d'essai amenant des conclusions différentes. Dans certains cas (1-2), il est ainsi possible de prendre une décision de qualification plus rapide permettant une sauvegarde de la dépense d'essai. D'autres cas, par

contre, nécessite une amélioration intermédiaire du plan à partir d'une nouvelle optimisation du plan d'essai.

Conclusions et Perspectives

Dans le monde d'aujourd'hui, la fiabilité prend une place de plus en plus importante et progresse dans tous les domaines industriels, de recherche, et également de la vie courante. Une fiabilité significative induit la possibilité d'un retour sur investissement dans un temps plus court pour une meilleure rentabilité, assurant une haute qualité de produits et surtout la possibilité d'atteindre les objectifs industriels, correspondant aux buts et aux intérêts des investisseurs, des entrepreneurs et des dirigeants d'entreprises ou d'organisations. Il est typique dans le développement d'une nouvelle technologie ou d'un système complexe d'avoir des objectifs de fiabilité. La fiabilité est une exigence obligatoire pour la satisfaction du client car il est possible qu'un produit non fiable nuise gravement à celle-ci, à contrario d'un produit fiable qui n'influe pas considérablement d'une manière positive. Si un produit ne fonctionne pas durant la période de garantie, les frais de remplacement et de réparation affecteront les profits, à l'inverse de l'attention négative qui lui sera portée. L'introduction des analyses de fiabilité est une étape importante dans la prise en compte de mesures correctives, conduisant finalement à un produit plus fiable.

Les industriels dépensent des sommes considérables afin d'améliorer la fiabilité de leurs produits chaque année avec beaucoup d'effort pour évaluer la fiabilité en organisant des collectes de données et des essais. La conception est souvent basée sur des données expérimentales de durée de vie, obtenues régulièrement à partir de seulement quelques unités. Pour certains produits, le temps d'essai nécessaire pour fournir avec assurance une estimation de la fiabilité dans des conditions normales de fonctionnement peut être excessivement long et très coûteux.

La collecte de données de fiabilité ne doit pas retarder le développement et doit être aussi économique que possible. Il est donc important d'être en mesure de procéder à des essais de fiabilité le plus rapidement possible, ceux-ci compatibles avec l'obtention de résultats significatifs.

Dans un essai classique, l'estimation de la fiabilité nécessite de tester de nombreux éléments pendant une longue période pour précipiter des défaillances; condition nécessaire à l'obtention d'une bonne estimation. Toutefois, ces conditions sont devenues incompatibles avec les contraintes industrielles actuelles.

Les essais de vie accélérée utilisent une variété de méthodes de tests avec contraintes élevées visant à raccourcir la durée de vie d'un produit ou accélérer la dégradation de sa performance. Le but est d'obtenir efficacement des données de performance qui, lorsqu'elles sont correctement analysées, permettent d'obtenir des estimations raisonnables en conditions normales de fonctionnement. Dans l'élaboration de ces essais, les ingénieurs sont souvent confrontés à deux problèmes :

- Quelle est la durée de fonctionnement équivalente selon la contrainte normale ?
- Les défaillances induites dans les conditions d'essais accélérés sont-elles les mêmes que celles qui pourraient se produire dans des conditions normales ?

Pour développer des modèles d'essais accélérés précis et légitimes répondant aux questions précédentes, le niveau de stress doit être proche ou sur la plage de fonctionnement normal. Des

extrapolations des résultats d'essai peuvent être alors effectués en utilisant des modèles empiriques de stress, plutôt que des modèles théoriques.

C'est dans ce contexte que les travaux de cette thèse se sont inscrits.

Dans le troisième chapitre, nous avons présenté les essais accélérés comme une alternative pour une estimation de la durée de vie, rapide et de bonne qualité, lorsque les produits à tester sont très fiables, conduisant à une censure importante si les essais se déroulent en conditions nominales. Une des limitations importante de ces essais consiste à ne pas changer le mode de défaillance par une sévérisation des conditions de test.

Le plan déterminé par régression sur la loi d'accélération a pour effet de propager estimation de l'incertitude obtenue en condition nominale. Par conséquent, nous avons proposé d'utiliser une approche bayésienne pour estimer les modèles standards de vie accélérée, en intégrant les informations disponibles sur la fiabilité du produit. L'objectif est d'obtenir des estimations de bonne qualité (intervalle de confiance "raisonnable") pour de petits échantillons. Dans le quatrième chapitre, la formalisation de cette approche d'estimation bayésienne des modèles SVA paramétrique est donc présentée. La qualité de l'estimation dépend principalement de la quantité de produits défaillants durant le test. Ainsi, la taille de l'échantillon, les niveaux de stress, les temps d'essai... sont des paramètres importants qui influencent fortement l'apparition de défaillances. Par conséquent, une attention particulière doit être portée sur le choix de ces paramètres en termes de statistiques industrielles.

En outre, nous avons proposé d'étudier la construction de plans optimaux dans le cinquième chapitre. Nous avons étendu notre approche pour développer une formulation théorique des paramètres suivant les différents degrés de liberté afin de définir un plan d'essai accéléré. Ce développement inclut plusieurs méthodes pour obtenir l'optimum, l'une utilisant les surfaces de réponses permettant d'avoir une vue globale de l'espace de solutions pour un plan d'essai comportant peu de paramètres, puis une méthode utilisant les algorithmes génétiques permettant d'obtenir l'optimum dans un cadre plus général. Les deux méthodes sont illustrés sur des exemples numériques, l'un permettant l'optimisation robuste d'un plan d'essai sur des roulements à billes, l'autre permettant la définition d'un plan d'essai accéléré optimal pour un module électronique de contrôle d'une pompe.

Dans le dernier chapitre, nous avons abordé la conduite d'un plan d'essai accéléré suivant une stratégie préalablement définie. En planifiant les essais dans un ordre judicieux, comme de commencer le test à la contrainte la plus élevée, il est possible d'obtenir rapidement des défaillances pour que nous puissions disposer d'informations préliminaires sur les paramètres. Ainsi, il est envisageable de tester le critère de qualification, pour prendre plus rapidement la décision, ou d'ajuster un plan d'essai en considérant ces premiers résultats. La démarche proposée s'appuie sur les facteurs de compatibilité entre les résultats d'essai (Vraisemblance) et les distributions a priori.

Plusieurs voies d'amélioration de la méthodologie peuvent être envisagées afin de la rendre plus efficace. La difficulté principale de la méthodologie proposée concerne les temps de calcul. Ceci s'explique par un recours important à la simulation de Monte Carlo. Ainsi, il est possible de remplacer l'estimation de l'espérance de la matrice de Fisher, par simulation de données d'essai, par la méthode des fonctions marginales ([Srivastava and Mittal, 2010]) permettant d'accélérer les temps de calcul. Il serait intéressant également d'étudier des cas d'optimisation plus complexes que ceux traités jusqu'à présent. En effet, l'intérêt des algorithmes génétiques est justement leur capacité à réaliser des optimisations en considérant des variables continues et discrètes. Ainsi, il

est possible de construire un plan optimal en recherchant le nombre de niveaux, nombre de lots, allocation de l'échantillon à chaque niveau, Ceci devrait faire l'objet d'une étude complémentaire en traitant plusieurs exemples simulés. Cela permettrait de valider le choix d'algorithmes génétiques utilisé en termes de performances. Il pourrait être intéressant également de tester d'autres formulations des coûts d'essai et d'exploitation et voir leurs impacts sur le résultat de l'optimisation. De plus, une méthodologie de conduite des essais est présentée dans le dernier chapitre. Elle est élaborée à partir d'une stratégie de conduite définie en ce sens. La formalisation des différentes stratégies permettrait de donner un cadre général à la conduite de plan d'essai accéléré.

D'autres plans d'essai accéléré pourraient être intéressants à optimiser comme ceux, par exemple, à contraintes échelonnées. En effet, ils présentent l'intérêt de pouvoir construire la métrique de fiabilité à partir d'un seul essai dès lors que l'on puisse mettre plusieurs produits en test en même temps. C'est souvent le cas en électronique où il est possible de mettre plusieurs unités dans une enceinte climatique, sur un pot vibrant, ... La formulation du problème d'optimisation et de robustesse proposée permet d'ouvrir la méthodologie à d'autres types de plans d'essai. En particulier, il serait intéressant d'étudier la construction de plans de dégradation. Plusieurs développements seront nécessaires pour l'intégration dans la méthodologie proposée. En particulier, il faudra traiter :

- Formulation des essais accélérés de dégradation pour les différents modèles d'analyse de données de dégradation (méthode des trajectoires inversibles, processus de Wiener, Processus Gamma,),
- Estimation bayésienne des paramètres de modèle de dégradation accélérée,
- Capitalisation de la connaissance sur les processus de dégradation pour la construction des distributions a priori,
- Validation sur des cas simulés.

Une dernière voie de développement concernerait le traitement de la robustesse. En effet, dans la thèse, la seule source de variabilité considérée porte sur les valeurs des paramètres du modèle SVA. Or, lorsque l'on souhaite construire un plan de qualification sur un nouveau produit, avec un degré d'innovation plus ou moins important, il est possible que la variabilité concerne également les choix de modèle. Aussi, il serait intéressant d'étudier la robustesse en intégrant la méconnaissance des modèles réels.

Bibliographie

- [Afnor, 1988] Afnor (1988). *Recueil de normes françaises : Fiabilité, Maintenabilité et Disponibilité*. Ed Afnor-Ute.
- [Arizono et al., 2008] Arizono, I., Kawamura, Y., and Takemoto, Y. (2008). Reliability tests for weibull distribution with variational shape parameter based on sudden death lifetime data. *Eur J Oper Res*, 189(2):570–574.
- [ASTE, 1993] ASTE (1993). *Le rôle des essais dans la maîtrise de la fiabilité*. ed ASTE.
- [Bagdonavicius et al., 2000] Bagdonavicius, V., Gerville-Réache, L., Nikoulina, V., and Nikoulin, M. (2000). Expériences accélérées: Analyse statistique du modèle standard de vie accélérée. *Statistique Appliquée*, XLVII(3):5–38.
- [Bagdonavicius and Nikulin, 1995a] Bagdonavicius, V. and Nikulin, M. (1995a). On accelerated testing of systems. *European Journal of Diagnosis and Safety in Automation*, vol. 5:307–316.
- [Bagdonavicius and Nikulin, 1995b] Bagdonavicius, V. and Nikulin, M. (1995b). *Semi-parametrics Models in Accelerated Life Testing*. Queen’s University.
- [Bagdonavicius and Nikulin, 1997a] Bagdonavicius, V. and Nikulin, M. (1997a). On application of step stresses in accelerated life testing. *C. R. Acad. Sci. Ser. I Math.*, 325(5):523–526.
- [Bagdonavicius and Nikulin, 1997b] Bagdonavicius, V. and Nikulin, M. (1997b). Transfer functional and semi-parametric regression. *Biometrika*, vol. 84:365 – 378.
- [Bagdonavicius and Nikulin, 2001] Bagdonavicius, V. and Nikulin, M. (2001). *Mathematical models in the theory of accelerated experiments*. World Scientific Publishing Co.
- [Bagdonavicius and Nikulin, 2002] Bagdonavicius, V. and Nikulin, M. (2002). *Accelerated life models*. Chapman&Hall/CRC.
- [Bayarri and Berger, 1998] Bayarri, M. and Berger, J. (1998). Quantifying surprise in the data and model verification (with discussion). *Bayesian Statistics*, 6:53–82.
- [Bellman, 1961] Bellman, R. (1961). *Adaptive Control Processes*.
- [Bousquet, 2006] Bousquet, N. (2006). Subjective bayesian statistics: agreement between prior and data. Technical report, Institut National De Recherche En Informatique Et En Automatique.
- [Bris, 2000] Bris, R. (2000). Bayes approach in rdt using accelerated and long-term life data. *Reliability Engineering and System Safety*.
- [Caruso and Dasgupta, 1998] Caruso, H. and Dasgupta, A. (1998). A fundamental overview of accelerated testing analytical models. In *IEEE Proceedings Annual Reliability and Maintainability Symposium*, USA.
- [Chapouille, 1972] Chapouille, P. (1972). *La Fiabilité*.
- [Clarotti, 1998a] Clarotti, C. (1998a). Les techniques fréquentielle et bayésienne au service le l’ingénieur de sûreté de fonctionnement. Technical Report Projet ISDF 08/96, ISDF.
- [Clarotti, 1998b] Clarotti, C. (1998b). Les techniques fréquentielle et bayésienne au service le l’ingénieur de sûreté de fonctionnement. Technical Report 08/96, ISDF.

- [Cocozza-Thivent, 1997] Cocozza-Thivent, C. (1997). *Processus stochastiques et fiabilité des systèmes*. Editeur Springer.
- [Congdon, 2001] Congdon, P. (2001). *Bayesian statistical modelling*. Ed Wiley.
- [Cross et al., 2007] Cross, R., Makeev, A., and Armanios, E. (2007). Simultaneous uncertainty quantification of fracture mechanics based life prediction model parameters. *Int J Fatigue*, 29(8):1510–1515.
- [Crowe and Feinberg, 2001a] Crowe, D. and Feinberg, A. (2001a). *Design for reliability*. Ed CRC Press.
- [Crowe and Feinberg, 2001b] Crowe, D. and Feinberg, A. (2001b). *Design for reliability*. Ed CRC Press.
- [Dorp and Mazzuchi, 2004] Dorp, J. R. and Mazzuchi, A. T. (2004). A general bayes exponential inference model for accelerated life testing. *Journal of statistical planning and inference*.
- [Dorp et al., 1996] Dorp, J. R. V., Mazzuchi, T., Fornell, G., and Pollock, L. (1996). A bayes approach to step-stress accelerated life testing. *IEEE transaction on reliability*, vol 45(3).
- [Dunson and Taylor, 2005] Dunson, D. and Taylor, J. (2005). Approximate bayesian inference for quantiles. *J. Nonparametric Stat.*, 17(3):385–400.
- [Finetti, 1974] Finetti, B. D. (1974). *Theory of probability: a critical introductory treatment*. Wiley.
- [Fisher, 1922] Fisher, R. (1922). On the mathematical foundations of theoretical statistics. *Philosophical Trans. of the Royal Society of London*, A 222:309–368.
- [Fournier, 1993] Fournier, J. (1993). *Fiabilité du logiciel*. Hermes, Paris.
- [Goldstein, 1999] Goldstein, M. (1999). Bayes linear analysis. *Encyclopaedia of statistical sciences*, 3:29–34.
- [Guerin et al., 2003] Guerin, F., Dumon, B., and Usureau, E. (2003). New methodology of weighting of prior distribution used bayesian approach applied to reliability estimation. In *Probability Method Conference*, Los Angeles, USA.
- [Guerin et al., 2004] Guerin, F., Dumon, B., and Usureau, E. (2004). *Compatibility between prior and likelihood distributions and weighting approach applied to reliability*. Progresses In Quality and Dependability. Editors, Bacivarov, Dumon, Wiele. A paraître.
- [Guerin et al., 2001] Guerin, F., Tebbi, O., and Dumon, B. (2001). Accelerated testing based on mechanical damage model. In *IEEE International Conference on Reliability and Maintainability*, Philadelphia, Pennsylvania, USA.
- [Hamada et al., 2001] Hamada, M., Martz, H. F., Reese, C. S., and Wilson, A. G. (2001). Finding near-optimal bayesian experimental designs via genetic algorithms. *American Statistician*, 55(3):175–181. WOS:000170299300001.
- [Han, 1997] Han, M. (1997). The structure of hierarchical prior distribution and its applications. *Chinese Operations Research and Management Science*, 6:31–40.
- [Hoang, 2003a] Hoang, P. (2003a). *Handbook of reliability engineering*. ed Springer.
- [Hoang, 2003b] Hoang, P. (2003b). *Handbook of reliability engineering*. ed Springer.
- [Holland, 1992] Holland, J. H. (1992). *Adaptation in natural and artificial systems*. MIT Press, Cambridge, MA, USA.

- [Ibrahim et al., 2001] Ibrahim, J., Chen, M., and Sinha, D. (2001). *Bayesian survival analysis*. Ed Springer.
- [Idée et al., 2001] Idée, E., Lannoy, A., and Meslin, T. (2001). Estimation of a lifetime law for equipment on the basis of a highly right multicensored sample and expert assesment.
- [kececioglu, 1944] kececioglu, D. (1944). *Reliability and life testing Handbook*. Ed Prentice Hall.
- [kececioglu and Sun, 1999] kececioglu, D. and Sun, F.-B. (1999). *Environmental stress screening : its quantification, optimization and management*. Ed Prentice Hall.
- [Lannoy, 1996] Lannoy, A. (1996). *Analyse quantitative et utilité du retour d'expérience pour la maintenance des matériels et la sécurité*. Eyrolles.
- [Lannoy and Procaccia, 1994a] Lannoy, A. and Procaccia, H. (1994a). *Méthodes avancées d'analyse des bases de données du retour d'expérience industriel*. Ed Eyrolles.
- [Lannoy and Procaccia, 1994b] Lannoy, A. and Procaccia, H. (1994b). *Méthodes avancées d'analyse des bases de données du retour d'expérience industriel*. Ed Eyrolles.
- [Lannoy and Procaccia, 2001] Lannoy, A. and Procaccia, H. (2001). *L'utilisation du jugement d'expert en sûreté de fonctionnement*. Ed. TEC&DOC.
- [Lantieri and Guérin, 2012] Lantieri, P. and Guérin, F. (2012). Optimization of a step-stress accelerated life test plan by genetic algorithm. *Quality assurance*, 58.
- [Laprie et al., 1995] Laprie, J. C., Arlat, J., Blanquart, J. P., Costes, A., Crouzet, Y., Deswarte, Y., Fabre, J. C., Guillermain, H., Kaâniche, M., Kanoun, K., Mazet, C., Powell, D., Rabéjac, C., and Thévenod, P. (1995). Guide de la sûreté de fonctionnement. Cépadués, Toulouse.
- [Levitin, 2006] Levitin, G. (2006). Genetic algorithms in reliability engineering. *Reliability Engineering & System Safety*, 91(9):975–976. WOS:000239648900001.
- [Ligeron and et M. Neff, 1984] Ligeron, J.-C. and et M. Neff, A. D. (1984). *La fiabilité en exploitation : Organisation et traitement des données*. Ed Techniques et Documentation-Lavoisier.
- [Lyonnet, 2006] Lyonnet, P. (2006). *Ingénierie de la fiabilité*.
- [Madi, 1997] Madi, T. M. (1997). Bayesian inference for partially accelerated life tests using gibbs sampling. *Microelectronics Reliability*, 37(8).
- [Manry et al., 1996] Manry, M., Hsieh, C.-H., Dawson, M., Fung, A., and S.J.Apollo (1996). Cramer rao maximum a-posteriori bounds on neural network training error for non-gaussian signals and parameters. *Internat. Journ. of Intelligent Control and Syst.*, 1:381–391.
- [Marseguerra et al., 2006] Marseguerra, M., Zio, E., and Martorell, S. (2006). Basics of genetic algorithms optimization for RAMS applications. *Reliability Engineering & System Safety*, 91(9):977–991. WOS:000239648900002.
- [Marts and Walter, 1982] Marts, H. and Walter, R. (1982). *Bayesian reliability analysis*. Ed John Willey& sons.
- [McLean, 2000] McLean, H. (2000). *HALT, HASS & HASA Explained : Accelerated Reliability Techniques*. ASQ Quality press.
- [Meeker and Escobar, 1986a] Meeker, W. and Escobar, L. (1986a). Algorithm as218 elements of the fisher information matrix for the smallest extreme value distribution and censored data. *Applied Statistic*, vol 65:80–86.
- [Meeker and Escobar, 1986b] Meeker, W. and Escobar, L. (1986b). Optimum accelerated life tests with type ii censored data. *J. of Statistical Computation and Simulation*, vol 23:273–297.

- [Meeker and Escobar, 1998] Meeker, W. and Escobar, L. (1998). *Statistical methods for reliability data*. Ed Wiley series in probability and statistics.
- [Mihalache, 2007] Mihalache, A. G. (2007). *MODELISATION ET EVALUATION DE LA FIABILITE DES SYSTEMES MECATRONIQUES : APPLICATION SUR SYSTEME EM-BARQUE*. PhD thesis, Ecole Doctorale d'Angers - Institut des Sciences et Techniques de l'Ingénieur d'Angers.
- [MIL-HDBK-338B, 1998] MIL-HDBK-338B (1998). *Military handbook electronic reliability design handbook*.
- [Mortureux, 2001] Mortureux, Y. (2001). La sûreté de fonctionnement : méthodes pour maîtriser les risques. *Techniques de l'Ingénieur*, AG4670:1–17.
- [Nelson, 1990a] Nelson, W. (1990a). *Accelerated Testing: Statistical Models, Test Plans and Data Analysis*. John Wiley & Sons.
- [Nelson, 1990b] Nelson, W. (1990b). *Accelerated Testing: Statistical Models, Test Plans and Data Analysis*. Ed. Wiley Series in Probability and Mathematical Statistics.
- [Nikulin et al., 2007] Nikulin, M., Gerville-Réache, L., and Couallier, V. (2007). *Statistique des essais accélérés*. Hermes Science.
- [O'Connor, 2003a] O'Connor, P. (2003a). Testing for reliability. *Quality and Reliability Engineering international*, vol 19:73–84.
- [O'Connor, 2003b] O'Connor, P. (2003b). Testing for reliability. *Quality and Reliability Engineering international*, vol 19:73–84.
- [O'Connor, 2002] O'Connor, P. D. T. (2002). *Practical Reliability Engineering Fourth Edition*. Wiley.
- [Pages and Gondran, 1980] Pages, A. and Gondran, M. (1980). *Fiabilité des systèmes*. Ed Eyrolles.
- [Pagès and Gondran, 1980] Pagès, A. and Gondran, M. (1980). *Fiabilité des systèmes*. Eyrolles, Paris.
- [Pham, 2003] Pham, H. (2003). *Reliability Engineering*. Springer-Verlang, New York, USA.
- [Poncet, 2005] Poncet, A. (2005). Fiabilité expérimentale : dimensionnement optimale d'essais sous contraintes. Master's thesis, UNIVERSITE JEAN MONNET.
- [Poncet et al., 2004] Poncet, A., Dumas, C., Schimmerling, P., Sarriau, O., and Miazga, C. (2004). Dimensionnement de chapeaux de paliers de moteurs turbo diesel par une approche fiabiliste. *Congrès SF2M, Paris*.
- [Procaccia, 1998] Procaccia, H. (1998). Fondements et applications des approches fréquentielle bayésiennes dans le domaine de la sûreté de fonctionnement. *Edition interne ISDF*.
- [Procaccia and Morilhat, 1996a] Procaccia, H. and Morilhat, P. (1996a). *Fiabilité des structures des installations industrielles*. Ed. Eyrolles.
- [Procaccia and Morilhat, 1996b] Procaccia, H. and Morilhat, P. (1996b). *Fiabilité des structures des installations industrielles*. Eyrolles, Paris.
- [Procaccia et al., 1992a] Procaccia, H., Piepszownik, L., and Clarotti, C. (1992a). *Fiabilité des équipements et théorie de la décision statistique fréquentielle et bayésienne*. Ed Eyrolles.
- [Procaccia et al., 1992b] Procaccia, H., Piepszownik, L., and Clarotti, C. (1992b). *Fiabilité des équipements et théorie de la décision statistique fréquentielle et bayésienne*. Ed Eyrolles.

- [Ringler, 1979a] Ringler, J. (1979a). Utilisation des techniques bayésiennes dans le cas de système complexe. *Revue de Statistique Appliquée*, vol XXVII n°2:–.
- [Ringler, 1979b] Ringler, J. (1979b). Utilisation des techniques bayésiennes dans le cas de système complexe. *Revue de Statistique Appliquée*, XXVII(2):–.
- [Robert, 1992] Robert, C. (1992). *L'analyse statistique bayésienne*. Edition ECONOMICA.
- [Robert, 2001] Robert, C. (2001). *The Bayesian Choice. From Decision-Theoretic Motivations to Computational Implementation (second édition)*. Springer-Verlag: New York.
- [Ross, 2003] Ross, T. D. (2003). Accurate confidence intervals for binomial proportion and poisson rate estimation. *Computers in biology and medicine*, vol 33:509–531.
- [Sander and Badoux, 1991a] Sander, P. and Badoux, R. (1991a). *Bayesian Methods in Reliability*. Kluwer Academic Publisher.
- [Sander and Badoux, 1991b] Sander, P. and Badoux, R. (1991b). *Bayesian Methods in Reliability*. Kluwer Academic Publisher.
- [Shyur et al., 1999a] Shyur, H.-J., Elsayed, E., and Luxhøj, J. (1999a). A general model for accelerated life testing with time-dependent covariates. *Nav Res Logist*, 46(2-3):303–321.
- [Shyur et al., 1999b] Shyur, H.-J., Elsayed, E. A., and Luxhøj, J. T. (1999b). A general model for accelerated life testing with time-dependent covariates. *Naval Research Logistics*, vol 46:303–321.
- [Somoreville et al., 1997] Somoreville, F. I., Dietrich, L. D., and Mazzuchi, A. T. (1997). Bayesian reliability analysis using the dirichlet prior distribution with emphasis on accelerated life testing run in random order. *Nonlinear analysis theory. Methods and applications*, 30.
- [Srivastava and Mittal, 2010] Srivastava, P. and Mittal, N. (2010). Optimum step-stress partially accelerated life tests for truncated logistic distribution with censoring. *Applied Mathematical Modelling*, 34:3166–3178.
- [Suhner, 1994] Suhner, M. (1994). *Utilisation de l'analyse bayésienne pour optimiser la démarche de Fiabilité*. Thèse, Université de Nancy.
- [Taboada, 2008] Taboada, H. A. (2008). Development of a multiple objective genetic algorithm for solving reliability design allocation problems. of the 2008 Industrial Engineering Research Conference.
- [Tebbi, 2005] Tebbi, O. (2005). *Estimation des lois de fiabilité en mécanique par les essais accélérés*. PhD thesis, Ecole Doctorale d'Angers - Institut des Sciences et Techniques de l'Ingénieur d'Angers.
- [Tiplica, 2007] Tiplica, T. (2007). Cours de statistique et probabilités.
- [Tribus, 1972] Tribus, M. (1972). *Décisions rationnelles dans l'incertain*.
- [Usureau, 2001] Usureau, E. (2001). Application des méthodes bayésiennes pour l'optimisation des coûts de développement des produits nouveaux. Master's thesis, Institut des Sciences et Techniques de l'Ingénieur d'Angers.
- [Usureau et al., 2004] Usureau, E., Guerin, F., and Dumon, B. (2004). *Démarche bayésienne et applications à la Sécurité de Fonctionnement*. éditions HERMES Lavoisier. Rédaction des chapitres Modélisation de l'a priori et compatibilité de l'a priori et de la vraisemblance.

- [Vassilious and Mettas, 2001] Vassilious, A. and Mettas, A. (2001). Understanding accelerated life-testing analysis. In *IEEE Proceedings Annual Reliability and Maintainability Symposium*, USA.
- [Villemeur, 1988a] Villemeur, A. (1988a). *Sûreté de fonctionnement des systèmes industriels : Fiabilité, Facteurs humains et Informatisation*. Collection EdF - Ed Eyrolles.
- [Villemeur, 1988b] Villemeur, A. (1988b). *Sûreté de fonctionnement des systèmes industriels: fiabilité, facteurs humains, informatisation*. Eyrolles.
- [Voiculescu and Guerin, 2005] Voiculescu, S. and Guerin, F. (2005). Bayesian estimation in accelerated life testing. *QUALITA*.
- [Voiculescu and Guérin, 2009] Voiculescu, S. and Guérin, F. (2009). *Fiabilité des systèmes embarqués*. Thèse doctorat, [s.n.], [S.l.].
- [Yang, 2007] Yang, G. (2007). *Life Cycle Reliability Engineering*. Wiley.
- [Yang and Yang, 1998] Yang, K. and Yang, G. (1998). Robust reliability design using environmental stress testing. *Quality and Reliability Engineering International*, vol 14:381–384.
- [Ye et al., 2010] Ye, Z., Li, Z., and Xie, M. (2010). Some improvements on adaptive genetic algorithms for reliability-related applications. *Reliability Engineering & System Safety*, 95(2):120–126. WOS:000272376900007.
- [Zhang and Meeker, 2006] Zhang, Y. and Meeker, W. Q. (2006). Bayesian methods for planning accelerated life tests. *Technometrics*, 48(1):49–60. ArticleType: research-article / Full publication date: Feb., 2006 / Copyright © 2006 American Statistical Association and American Society for Quality.

Planification des essais accélérés : optimisation, robustesse et analyse

Seyyede Zohreh FATEMI

Résumé

La qualification d'un produit, lors des phases de développement, est une étape importante dans un projet qui vérifie que les performances et la fiabilité atteignent les objectifs. Les essais de qualification sont souvent coûteux en temps et en nombre de produits testés. Les essais accélérés consistent à soumettre des unités à des niveaux de stress plus élevés qu'en condition d'exploitation afin de réduire le temps d'apparition des défaillances. Ils permettent de construire plus rapidement la fonction de fiabilité à partir d'un modèle appropriée reliant durée de vie et stress. De plus, un plan d'essai doit être construit, précisant les paramètres du plan (niveaux de stress, allocation de l'échantillon) pour trouver le meilleur compromis entre le coût d'essai et la qualité d'estimation. L'objectif de la thèse a été de définir une méthodologie de construction de plans d'essai accéléré optimaux et robustes. Nous avons donc développé un cadre général basé sur la minimisation du coût global et une approche bayésienne. Les distributions a priori, issues de la connaissance des paramètres de fiabilité et du modèle d'accélération, sont utilisées dans l'inférence bayésienne et dans une simulation de Monte Carlo d'exploration des fiabilités possibles. Le plan optimal et robuste est obtenu à partir de méthodes d'optimisation (Surface de réponse, Algorithmes génétiques). Enfin, une méthodologie de suivi des essais est développée en observant la pertinence des résultats par rapport aux informations a priori à partir d'un facteur de similitude. Il permet de vérifier si la décision quant à la qualification peut être prise plus rapidement, ou d'optimiser le plan en cours de réalisation.

Mots-clés : Fiabilité, Essais accélérés, optimisation, estimation bayésienne, coût des essais, coût d'exploitation, robustesse

Abstract

Qualification of a product during the development phases is an important step in a project. It permits to verify that the performances and reliability achieve the objectives. The qualification tests are often costly in time and number of products tested. Accelerated tests consist of submitting units to stress levels higher than operating condition to reduce the time to failures and allow building faster the reliability function. This requires the determination or development of appropriate life-stress relationship model. In addition, a test plan should be built, accurate the design parameters (number of test levels, stress levels, sample allocation at each stress level ...) to find the best compromise between cost test and quality estimation. The objective of the thesis is to define a methodology to obtain optimal and robust accelerated test plans. So, we have developed a general framework based on the minimization of a global cost function and a Bayesian approach. The prior distributions, from the available knowledge on the parameters of reliability and acceleration model, are used in Bayesian inference and in a Monte Carlo simulation to explore possible reliabilities. The optimal and robust plan is obtained by optimization methods (Response surface and Genetic algorithms). Finally, a methodology of monitoring during test realization is developed. The relevance of the results in relation to prior information is studied using a similarity factor between observed and a priori data. It allows verifying whether the decision on the qualification can be taken more quickly in presence of more reliable product than expected or optimizing the plan under realization.

Keywords: accelerated test, reliability, optimisation, Bayesian estimation, testing cost, operation cost, robustness