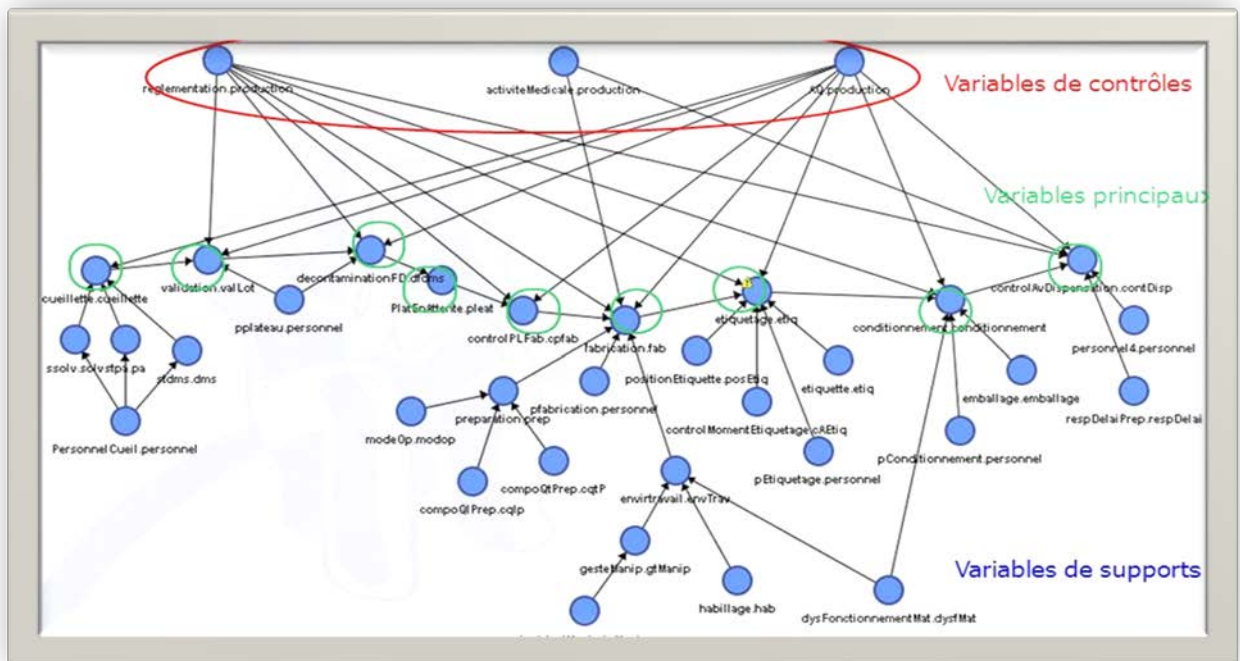


الشعبية الديمقراطية الجزائرية الجمهورية
العلمي والبحث العالي التعليم وزارة
République Algérienne Démocratique et Populaire
Ministère de l'Enseignement Supérieur et de la Recherche Scientifique
حمد أبن محمد 2 وهران جامعة
الصناعي الأمن و الصيانة معهد
Université d'Oran 2 Mohamed Ben Ahmed
Institut de Maintenance et de Sécurité Industrielle

Démarche Bayésienne pour la sûreté de fonctionnement (Support de Cours)



Pour les classes : 2ème année Master : Sécurité Industrielle et environnement

Présenté par : Dr AISSANI Nassima

Sommaire

Sommaire	2
Liste des figures.....	5
Liste des tableaux.....	6
Préambule.....	7
CHAPITRE 1 : Notions fondamentales de la sureté de fonctionnement	8
1. Introduction.....	8
2. Définitions	8
2.1 Défaillance.....	8
2.2 Défaut.....	9
2.3 Panne	9
3. Modes, causes et mécanismes de défaillance	10
4. Etude du système en vue d'une analyse de sureté de fonctionnement	10
4.1 L'étude fonctionnelle	11
4.2 Description structurale des procédés industriels	13
CHAPITRE 2 : Introduction à la fiabilité et la sureté de fonctionnement	16
1. La fiabilité	16
1.1 Sélection d'un échantillon	16
1.2 Formalisme :.....	17
1.3 Les mesures de la fiabilité	18
1.4 Densité de défaillance	19
1.5 Taux de défaillance :.....	19
1.6 Application exemple de l'étude de la fiabilité	20
2. La disponibilité.....	27
2.1 Les grandeurs moyennes associées à la disponibilité :.....	27
2.2 Disponibilité opérationnelle	28
3. La maintenabilité	28
3.1 Mesures liées à la maintenabilité	28
3.2 Taux de réparation $\mu(t)$	29
4. Analyse de fiabilité, maintenabilité, disponibilité possibles	29

5. La sécurité	29
CHAPITRE 3 : Les modèles de SDF: Les modèles avancés	31
1. Application de l'approche classique de probabilité dans la sureté de fonctionnement ..	32
2. Application de l'approche Fréquentiste dans la sureté de fonctionnement.....	36
2.1 Définition de la fonction de répartition.....	37
2.2 Relations entre $R(t)$, $F(t)$, $f(t)$ et $\lambda(t)$	37
3. Application de l'approche subjective dans la sureté de fonctionnement.....	39
CHAPITRE 4 : Les bases de l'approche bayésienne	41
1. Introduction.....	41
2. La probabilité conditionnelle	41
2.1 Espérance conditionnelle	42
2.2 Densité conditionnelle.....	42
2.3 Représentation par un arbre pondéré.....	42
3. Généralisation du théorème de Bayes.....	45
4. Formulation continu du théorème de Bayes	47
5. L'approche bayésienne adoptée en pratique	48
6. Le mécanisme d'inférence	50
CHAPITRE 5 : Réseaux bayésiens et l'outil« BayesiaLab ».....	51
1. Introduction :.....	51
2. Structure du Réseau Bayésien (RB).....	51
2.1 Tables de probabilités	53
2.2 Notion de causalité et relations entre les nœuds.....	53
2.3 Cas de deux variables discrètes	53
2.4 Cas de deux variables discrète et continue	54
2.5 Cas de deux variables continues.....	54
3. Calculs d'inférences dans un Réseau Bayésien	54
4. Exemple d'un réseau bayésien	55
5. Structure du réseau bayésien à partir de l'arbre des défaillances " conditions pour une mutation de l'AdD au RB"	57
6. Introduction sur BayesiaLab	59
7. Modélisation de la connaissance par Bayesialab.....	59

8. Exemple d'une simulation avec BayesiaLab.....	60
CHAPITRE 6 : Modélisation des données et principe de conjugaison	65
1. Introduction.....	65
2. Principe de conjugaison.....	66
3. Fonction de vraisemblance Exponentielle	67
3.1 Distribution a priori Gamma.....	67
3.2 Fonction de vraisemblance.....	67
3.3 Densité a posteriori	68
CHAPITRE 7 : Actualisation Bayésienne	72
1. Introduction.....	72
2. Les conditions et contraintes de l'actualisation.....	72
3. Actualisation des paramètres de fiabilité	72
4. La méthode des moments	73
Références (Par ordre d'apparition).....	75

Liste des figures

Figure 1 Apparition de la défaillance	9
Figure 2 Schéma d'une description fonctionnelle.....	13
Figure 3 Dynamo	15
Figure 4 Description matériel de la dynamo	15
Figure 5 Evolution de la densité de fiabilité.....	17
Figure 6 Taux de défaillance par intervalle de temps	18
Figure 7 Courbe en baignoire	20
Figure 8 Schéma de la machine de découpe	21
Figure 9 Représentation graphique du taux de défaillance en fonction du temps	22
Figure 10 Représentation graphique du taux de Fiabilité en fonction du temps.....	23
Figure 11 Dates du changement systématique des supports des lames de découpe.....	26
Figure 12 Evolution de la probabilité de disponibilité	27
Figure 13 Les mesures de disponibilité.....	28
Figure 14 Les durées caractéristiques de Fiabilité-Maintenabilité-Disponibilité	28
Figure 15 Exemple d'un arbre de défaillance de l'alimentation électrique	35
Figure 16 Allure de R(t) loi exponentielle	37
Figure 17 Allure de R(t) loi Weibull.....	39
Figure 18 Thomas Bayes	41
Figure 19 Arbre pondérée	42
Figure 20 Imbrication des événements	45
Figure 21 Probabilité a priori et a posteriori.....	46
Figure 22 Variable continue.....	47
Figure 23 Les données à priori.....	49
Figure 24 Mécanisme d'inférence	50
Figure 25 Arbre de probabilité	56
Figure 26 Réseau bayésien par BayesiaLab 1.....	56
Figure 27 Réseau bayésien par BayesiaLab 2.....	56
Figure 28 Réseau bayésien par BayesiaLab (Il a plu pendant la nuit).....	57
Figure 29 Transformation depuis l'arbre des défaillances au réseau bayésien.....	58
Figure 30 Figure 1. Présentation des portes logiques " OU et ET " dans l'AdD et le RB.....	58
Figure 31 Création des nœuds.....	60
Figure 32 Propriété du nœud maladie.....	61
Figure 34 Les deux modes	62
Figure 33 Table des probabilités du nœud test	62
Figure 35 Les modalités	63
Figure 36 Changement des modalités	63
Figure 37 Principe d'actualisation.....	72

Liste des tableaux

Tableau 1 Statistiques des défaillances	18
Tableau 2 Calcul du taux de défaillance exprimé en défaillance / heure	21
Tableau 3 <i>Calcul de la fonction fiabilité R(t)</i>	23
Tableau 4 Ce tableau indique les analyses correspondant au type de système [5].	29
Tableau 5 <i>Recueils de données de fiabilité en électronique [11]</i>	33
Tableau 6 relations entre $F(t)$, $R(t)$, $f(t)$ et $I(t)$	38
Tableau 7 lois exponentielle et de Weibull	38
Tableau 8 table de probabilités conditionnelles à deux variables discrètes	54
Tableau 9 table de probabilité à variables discrètes et continues	54
Tableau 10 Probabilité a posteriori	55
Tableau 11 Principales lois conjuguées.....	66

Préambule

La Sûreté de fonctionnement s'est développée principalement au cours du 20e siècle pour être actuellement un domaine incontournable pour les industries à risques mais aussi, de plus en plus, pour toute l'industrie, en raison de sa corrélation avec la notion de qualité et l'impact sur l'environnement. Jusqu'avant la Seconde Guerre mondiale, la Sûreté de fonctionnement est demeurée plus un art empirique qu'une science exacte. Mais au cours de cette dernière décennie, des méthodes plus analytiques voient le jour pour prendre des décisions mieux adaptées au cas étudiés. Nous pouvons citer : AMDE et AMDEC, Arbre des Défaillances, HAZOP, Arbre des événements, Noeud papillon... Ces méthodes ont pour but d'identifier les risques dans un procédé industriel et quantifier leur criticité, gravité et probabilité d'apparition, elles permettent aussi de réaliser une cartographie des risques. La cartographie des risques reste l'outil fédérateur pour le pilotage et la prise de décision, cette cartographie sert à recenser tous les risques, les hiérarchiser et fournir une vision d'ensemble aux décideurs afin d'orienter les stratégies d'actions.

Suite à ces études analytiques, des études de quantifications des scénarios de risque, peuvent être réalisées afin d'optimiser leur coût, citant : les réseaux bayésien et la méthode monte carlo. Dans ce cours nous nous intéresserons surtout à l'approche bayésienne.

L'approche bayésienne permet d'introduire de la variabilité dans les scénarios étudiés et de simuler les pertes sans nécessiter d'une base historique. Elle permet également d'identifier des leviers de réduction de risque dans la mesure où les pertes sont conditionnées à des indicateurs qui peuvent être suivis et modifiés pour réduire le risque. L'approche bayésienne a pour avantage d'être formelle basée sur la théorie des probabilités ce qui permet de réaliser une modélisation dont il est facile d'analyser en lançant des scénarios de simulation. Enfin, la réalisation d'un réseau bayésien pour un cas d'étude permet d'augmenter le système dynamique de gestion des risques par un module d'aide à la décision basé sur les réseaux bayésiens.

CHAPITRE 1 : Notions fondamentales de la sûreté de fonctionnement

1. Introduction

Dans un contexte international extrêmement concurrentiel, les entreprises doivent maîtriser les différents outils qui leur permettent de rester compétitives et doivent s'engager dans des actions d'amélioration à tous les niveaux. La complexité croissante des systèmes, la réduction de leurs coûts de conception et d'exploitation, leur utilisation de plus en plus importante dans la vie quotidienne font que la sûreté de fonctionnement est devenue incontournable dans le développement de tout système industriel.

La sûreté de fonctionnement (SdF) fait partie des enjeux majeurs de ces dernières années et des années à venir. Cette notion désigne à la fois un ensemble de moyens et un ensemble de résultats produits par ces moyens:

- des méthodes et des outils pour caractériser et maîtriser les effets des aléas, des pannes et des erreurs ;
- la quantification des caractéristiques des systèmes pour exprimer la conformité dans le temps de leurs comportements et de leurs actions.

2. Définitions

Dans un premier temps, des définitions seront posées sur des notions fondamentales, telles que les défaillances, pannes.... [1]

2.1 Défaillance

Une défaillance est « l'altération ou la cessation de l'aptitude d'un ensemble à accomplir sa ou ses fonction(s) requise(s) avec les performances définies dans les spécifications techniques ». d'après la norme AFNOR 60010X

L'ensemble est indisponible suite à la défaillance.

La cessation de l'aptitude conduit l'entité à être dans un état appelé panne.

Exemple : si l'on considère un moteur électrique dont la fonction principale est de convertir une énergie électrique en énergie mécanique, le refus de démarrage est une défaillance fonctionnelle du moteur.

Il existe plusieurs types de défaillances.

- Progressive: Défaillance due à une évolution dans le temps des caractéristiques d'une entité. En général, une telle défaillance peut être prévue par un examen ou une surveillance antérieure

- Soudaine: Défaillance qui ne se manifeste pas par une perte progressive des performances et qui ne pouvait être prévue par un examen ou une surveillance.
- Partielle: Défaillance résultant d'une déviation d'une des caractéristiques au-delà des limites spécifiées, mais telle qu'elle n'entraîne pas une disparition complète de la fonction requise.
- Complète: Défaillance résultant de déviation d'une ou des caractéristiques au-delà des limites spécifiées telle qu'elle entraîne une disparition complète de la fonction requise.
- Catalectique: Défaillance à la fois soudaine et complète
- Par dégradation: Défaillance qui est à la fois progressive et partielle. Avec le temps, une telle défaillance peut devenir une défaillance complète.

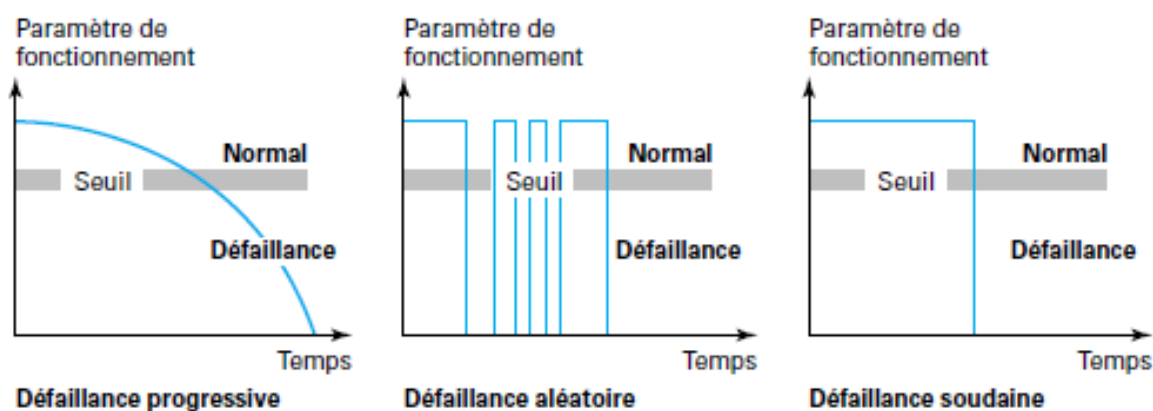


Figure 1 Apparition de la défaillance

2.2 Défaut

Un défaut est un écart dépassant les limites acceptables dans des conditions données entre la caractéristique d'une entité et la caractéristique voulue. Un défaut peut mais ne conduit pas nécessairement à une défaillance et un défaut d'un composant ou d'un système peut ne pas affecter l'aptitude d'un système à accomplir une fonction requise.

2.3 Panne

La panne est l'inaptitude d'une entité à accomplir une fonction requise. Une panne résulte toujours d'une défaillance. Les pannes se classent en quatre catégories:

- Panne intermittente: Panne d'une entité subsistant pendant une durée limitée après laquelle l'entité redevient apte à accomplir une fonction requise sans avoir été soumise à une action de maintenance corrective.
- Panne fugitive: Panne d'une entité qui est intermittente et difficile à constater.
- Panne permanente: Panne d'une entité qui persiste tant que des opérations de maintenance corrective n'ont pas eu lieu.
- Panne latente: Panne qui existe mais qui n'a pas encore été détecté.

3. Modes, causes et mécanismes de défaillance

Un mode de défaillance est la manifestation extérieure d'un phénomène physique anormal, par laquelle on constate la détérioration ou la perte d'au moins une des caractéristiques de fonctionnement d'une entité.

Une défaillance résulte de causes qui se définissent comme des circonstances liées à la conception, la fabrication ou l'emploi. Ces causes se manifestent par l'intermédiaire d'un mécanisme de défaillance qui se définit comme tout processus physique, chimique ou autre qui entraîne une défaillance.

Un mode de défaillance modifie l'état du composant, i.e. l'ensemble des paramètres caractérisant le fonctionnement de ce composant. Pour chaque type de matériels, un certain nombre de modes de défaillance est déterminé à l'avance.

Dans une étude de fiabilité de système, on s'intéressera aux modes de défaillance d'un composant qui sont supposés avoir un impact sur le fonctionnement du système auquel il est connecté.

4. Etude du système en vue d'une analyse de sûreté de fonctionnement

L'objet sous étude est le système et les fonctions qu'il fournit. Il existe de nombreuses définitions de système dans le domaine des systèmes d'ingénierie.

Définition : Un système peut être décrit comme un ensemble d'éléments en interaction entre eux et avec l'environnement dont le comportement dépend : des comportements individuels des éléments qui le composent, des règles d'interaction entre éléments (interfaces, algorithmes, protocoles), de l'organisation topologique des éléments (architectures).[2]

Une étude de Sûreté de fonctionnement passe nécessairement par une analyse exhaustive des différentes phases de fonctionnement faisant appel à des termes précis tels que missions et fonctions assurées par un bien. Il est particulièrement indispensable de définir les notions de fonction et de mission. Ces définitions sont, en effet, très variables d'un secteur d'activité à un autre ou d'un pays à un autre. Les missions et fonctions d'une entité découlent directement de la définition et de la spécification de ses exigences opérationnelles:

- quelles sont les exigences opérationnelles et comment l'entité sera-t-elle utilisée ?
- où l'entité sera-t-elle utilisée?
- combien de temps l'entité sera-t-elle utilisée?

4.1 L'étude fonctionnelle

4.1.1 La mission

La mission est décomposée en un ensemble de phases qui sont représentés par des fonctions. La mission est considérée comme satisfaite si toutes les phases sont accomplies.

4.1.2 La Fonction

La compréhension de la notion de fonction et de ses paramètres est l'élément clé sur lequel reposent les analyses de Sécurité de fonctionnement car on recherche les éléments matériels qui vont entraîner la perte ou la dégradation des fonctions.

L'AFNOR définit une fonction comme « l'action d'une entité ou de l'un de ses constituants exprimée en termes de finalité ».

Cette définition de nature qualitative est indispensable pour décrire de façon globale une fonction. La fonction fait appel à des notions qualitatives et quantitatives.

Pour un système tel qu'un propulseur à poudre, la description de sa fonction est simple : fournir une poussée nominale de 15 t pendant 30 s. Cette définition donne ses caractéristiques en termes quantitatifs. Un moteur électrique possède une fonction de base consistant à convertir de l'énergie mécanique en énergie électrique.

Les différences entre ces fonctions sont quelquefois subtiles et l'acception de leur terminologie doit toujours faire l'objet d'une acceptation au sens d'une même entreprise et de ses prestataires.

Pour des systèmes plus complexes, il est indispensable de classer et de hiérarchiser la nature des fonctions :

- principales ;
- secondaires ;
- de protection ;
- redondantes.

4.1.3 Fonctions principales

Une fonction principale peut se définir comme étant la raison d'être d'un bien ou d'un système défini souvent avec ses caractéristiques associées (durée, caractéristiques physiques, chimiques...).

Par exemple, une première définition générale de la fonction principale d'une chaudière est de fournir de la vapeur.

Une seconde définition de la fonction de la chaudière à vapeur est de fournir de la vapeur à 110 °C avec un débit de 40 t/heure pendant 24 h.

Ces deux exemples de définitions montrent la difficulté de décrire une fonction quand on s'intéresse à ses modes de défaillance. En effet, dans la première définition, la perte de la fonction correspondra à une non-fourniture de vapeur. Il n'existera dans ce cas qu'un seul mode de défaillance.

4.1.4 Fonctions secondaires

Dans de nombreux cas, un système assure d'autres fonctions que la fonction principale. Ces fonctions sont appelées fonctions secondaires et leur perte peut également avoir des conséquences catastrophiques.

En reprenant l'exemple de la chaudière, une fonction secondaire est de maintenir l'intégrité du confinement de la vapeur. L'existence d'une fuite ou le risque d'une explosion entraînerait une défaillance de la fonction principale. Le calorifugeage de la chaudière est une autre fonction secondaire de la chaudière ayant pour but de minimiser les pertes thermiques.

4.1.5 Fonctions de protection

Les fonctions de protection ont pour but de garantir, par des moyens de signalisation ou la mise en route de systèmes redondants, la sécurité des biens, des personnes et de l'environnement. Ces fonctions de protection sont assurées par des systèmes de signalisation, d'alarme ou de protection automatique.

Dans le cas d'une chaudière à vapeur alimentée par un brûleur à gaz, les soupapes de sécurité assurent une protection passive contre le risque de surpression, le système de mesure de pression et de température permet aux opérateurs de contrôler les anomalies de fonctionnement et le système de détection de gaz à l'intérieur de la chaudière a pour rôle d'éviter son explosion.

4.1.6 Fonctions redondantes

Dans les industries telles que celles des secteurs aéronautiques, nucléaires et spatiaux, des systèmes ou des matériels redondants (doublés, triplés ou quadruplés) sont couramment mis en œuvre pour assurer le niveau requis de sécurité ou de sûreté. Ces systèmes redondants peuvent fonctionner en permanence (redondance active) ou être en attente (redondance passive). Dans l'industrie automobile, c'est ainsi que l'on équipe certains véhicules avec un double circuit de freinage. Pour un avion bimoteur, les deux moteurs illustrent la redondance active. En effet, en cas de panne d'un des moteurs, le propulseur restant a été conçu pour pouvoir ramener l'aéronef sur un terrain d'atterrissage avec toutes les conditions de sécurité.

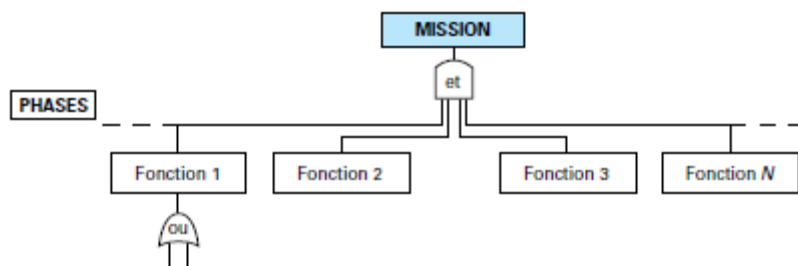


Figure 2 Schéma d'une description fonctionnelle

4.2 Description structurelle des procédés industriels

Pour assurer ces objectifs fonctionnels de haut niveau, le processus fait appel à un ensemble de systèmes interconnectés ou en interaction. Chaque système assure une ou plusieurs fonctions bien définies (s).

Dans la description structurelle on s'intéresse à la morphologie physique du système: fournit essentiellement tous les éléments constitutifs sans se préoccuper de leurs fonctions.

La notion de processus industriel recouvre des secteurs industriels très variés. Au sens très large, un processus assure la fabrication d'un produit ou fournit un service. Un avion assure un service de transport de biens ou de personnes. Un ordinateur s'assimile également à un processus puisqu'il fournit un service de calculs scientifiques ou de gestion. Ces exemples illustrent la notion très large de processus industriels. Conceptuellement, on appellera processus industriel, une installation complexe assumant un objectif fonctionnel de haut niveau (production de biens ou de services).

Pour assurer ces objectifs fonctionnels de haut niveau, le processus fait appel à un ensemble de systèmes interconnectés ou en interaction. Chaque système assure une ou plusieurs fonctions bien définie(s). Les systèmes peuvent être décomposés en sous-systèmes assumant à leur tour généralement un seul objectif fonctionnel.

4.2.1 Description générale

Les termes suivants s'emploient pour décrire les entités constitutives d'un processus ou d'une installation industrielle sans vouloir leur affecter un sens matériel ou fonctionnel.

4.2.1.1 Bien durable

Tout élément, composant, équipement, sous-système, système, matériel de processus, etc., que l'on peut considérer individuellement et qui a pour objectif d'assurer une fonction donnée pendant un temps relativement long, compte tenu de la qualité des opérations de maintenance. Un bien durable peut être relativement simple (robot de cuisine) ou complexe (avion, centrale nucléaire, ouvrage d'art, etc.).

4.2.1.2 Élément

Partie constitutive d'un ensemble ou sous-ensemble quelles qu'en soient la nature ou la dimension.

4.2.1.3 Sous-ensemble

Groupement d'éléments associés en fonctionnement entrant dans la constitution d'un ensemble. Le sous-ensemble peut avoir une signification matérielle et/ou fonctionnelle.

4.2.1.4 Ensemble

Groupement de sous-ensembles assurant une ou plusieurs fonctions techniques qui le rendent apte à remplir une fonction opérationnelle. L'ensemble peut avoir également une signification matérielle et/ou fonctionnelle.

4.2.2 Description fonctionnelle

Une description fonctionnelle se présente généralement sous la forme d'une arborescence hiérarchisée à plusieurs niveaux. Cette décomposition fonctionnelle sous forme d'arbre fonctionnel n'est pas la seule méthode disponible. Des méthodes issues des techniques de l'analyse de la valeur et de l'analyse fonctionnelle sont utilisables pour décrire les phases de conception et l'exploitation d'un système industriel (méthodes FAST, RELIASEP®, APTE®, SADT®...). En théorie, cette description fonctionnelle devrait faire abstraction de toute réalisation matérielle. Dans la pratique, surtout pour la description d'installations complexes opérationnelles, on associe très souvent les éléments matériels qui contribuent à réaliser ces fonctions. Trois termes principaux sont largement utilisés pour décrire fonctionnellement une installation industrielle complexe : les systèmes, les sous-systèmes et les composants.

4.2.2.1 Système

C'est l'association de sous-systèmes constituant un tout organique complexe, destiné à remplir une fonction générale d'un bien durable complexe.

4.2.2.2 Sous-système

Le sous-système représente une association de composants destinée à remplir une ou plusieurs fonctions opérationnelles.

4.2.2.3 Composants

Le composant représente un élément matériel ou un ensemble matériel remplissant une fonction particulière dans un système ou sous-système. Si l'installation industrielle est complexe il est possible de compléter la description par des niveaux inférieurs tels que sous-sous-systèmes, etc.

4.2.3 Description matérielle

Une description matérielle d'un processus ou d'un équipement fournit essentiellement tous ses éléments constitutifs sans se préoccuper de leurs fonctions. On peut considérer cette décomposition comme une mise à plat des éléments que l'on obtiendrait en effectuant un démontage complet. Elle correspond à une logique de démontage / remontage. De ce fait, elle donne une représentation externe et physique. La figure 4 représente la décomposition matérielle de la machine à laver la vaisselle. Dans l'industrie, cette méthode de décomposition est particulièrement utile pour établir une nomenclature des différents éléments pour la gestion des stocks, la mise en œuvre d'un outil de gestion de maintenance assistée par

ordinateur (GMAO) et, dans certains cas, pour servir de support aux banques de données de retour d'expérience. Cette décomposition matérielle comprend des descripteurs matériels tels que pièces, organes, mécanismes, dispositifs, matériels et installations qui vont du plus petit élément (pièce) au plus important (installation).

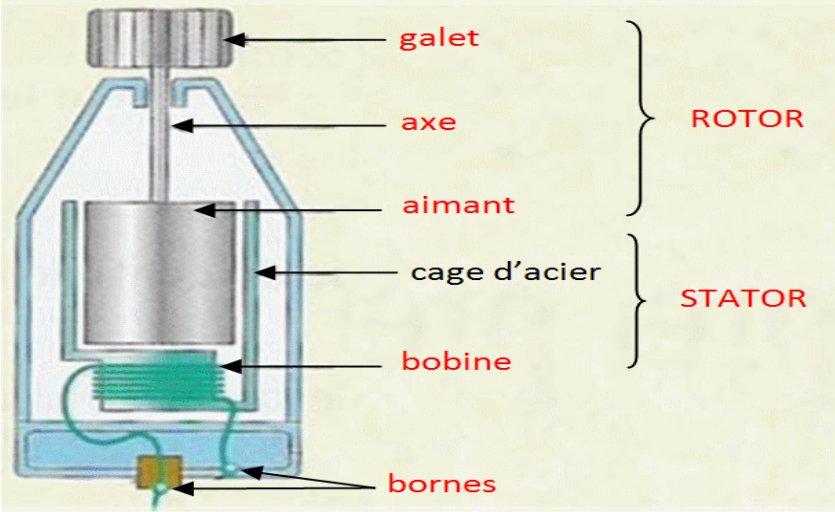


Figure 3 Dynamo

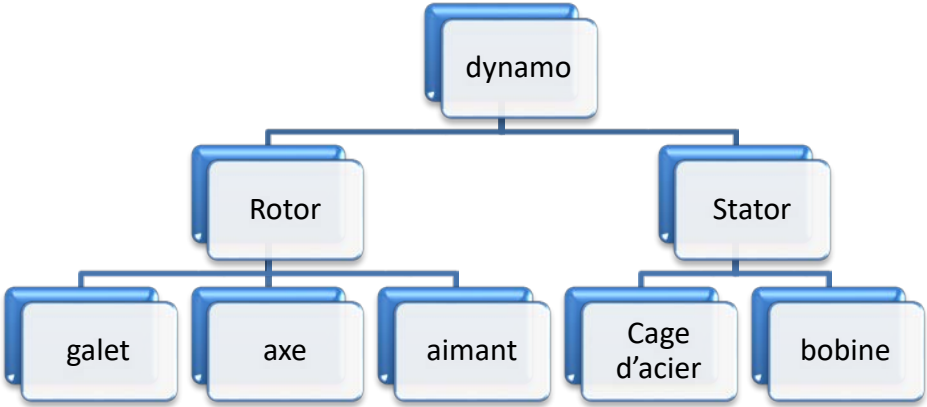


Figure 4 Description matériel de la dynamo

CHAPITRE 2 : Introduction à la fiabilité et la sûreté de fonctionnement

La Sûreté de fonctionnement est appelée la science des « défaillances ». Appelée aussi : analyse de risque (milieu pétrolier), cyndinique (science du danger), FMDS (Fiabilité, Maintenabilité, Disponibilité, Sécurité), en anglais RAMS (Reliability, Availability, Maintainability and Safety). [3]

Elle se caractérise à la fois par les études structurelles statiques et dynamiques des systèmes, du point de vue prévisionnel mais aussi opérationnel et expérimental (essais, accidents), en tenant compte des aspects probabilités et des conséquences induites par les défaillances techniques et humaines [4].

La sûreté de fonctionnement est basée sur 4 notions fondamentales [5]:

1. La fiabilité

Est la probabilité de non-défaillance d'un équipement sur un intervalle de temps. Introduite en 1962 pour traduire le terme anglais *reliability*, sert à étudier l'aptitude de systèmes à fonctionner correctement durant une période donnée. Un dispositif peut se trouver dans l'un des deux états suivants :

1. Apte à fonctionner correctement, c'est-à-dire en état de service.
2. Inapte à fonctionner correctement, c'est-à-dire en panne ou hors-service

Deux cas se présentent :

- Pour les grands échantillons ($N > 50$), il faut mettre en œuvre une estimation statistique en application de la statistique descriptive. Cette méthode donnera des estimateurs empiriques de la fiabilité de l'échantillon.
- Pour des échantillons petits ($8 < N < 50$), il faut mettre en œuvre une recherche de loi de probabilité ajustable à l'échantillon. Cette méthode fournira les caractéristiques de la population mère associée à un intervalle de confiance, sous forme graphique et/ou analytique.

1.1 Sélection d'un échantillon

L'étape de sélection d'un échantillon est importante, car il est essentiel d'obtenir une population homogène pour laquelle il sera possible de calculer des caractéristiques de fiabilité. Les principaux critères de sélection sont:

- les dates de début et de fin d'observation,
- les équipements concernés (le type, la famille et l'état des équipements, le niveau de décomposition dans l'arborescence matérielle, les limites des équipements),

- éventuellement un mode de défaillance pour ces équipements, si l'on s'intéresse à une évaluation de sécurité, ou à un calcul de disponibilité,
- les caractéristiques techniques: l'échantillon doit comprendre des équipements ayant les mêmes caractéristiques de conception,
- les conditions de fonctionnement: le type d'utilisation doit être similaire, les temps de fonctionnement (heures de fonctionnement et sollicitations) doivent être récupérés.
- les conditions environnementales (température, humidité, corrosion, ...).

1.2 Formalisme :

La fiabilité à la date t , notée $R(t)$: probabilité d'atteindre la date t sans défaillance.

Sachant que :

- Entité (**E**) désigne: composant/système/ sous-système...
- **T** l'instant d'occurrence de la défaillance

$R(t) = P(E \text{ non défaillante sur la durée } [0, t], \text{ en supposant qu'elle n'est pas défaillante à l'instant } t = 0).$

$$R(t) = P(T > t)$$

La défaillance à la date t (notée $F(t)$) : probabilité de défaillance avant t .

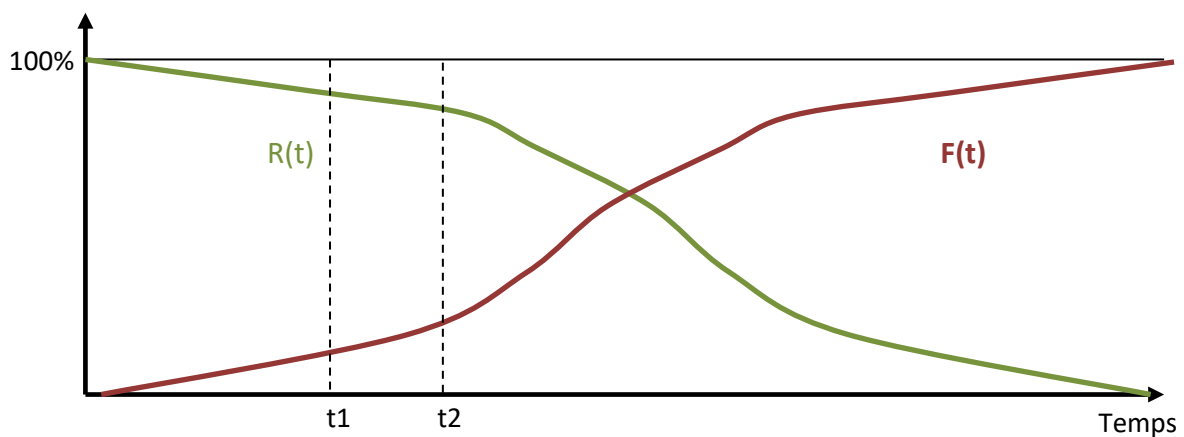


Figure 5 Evolution de la densité de fiabilité $R(t)$ et défaillance $A(t)$

entre $t1$ et $t2$ est donnée par l'expression :

$$P_a = (1 - R(t2)) - (1 - R(t1)) = R(t1) - R(t2) \text{ si } t2 > t1$$

La probabilité P_b que l'entité ait une défaillance entre $t1$ et $t2$, en supposant le système non défaillant à $t1$ est donnée par :

$$P_b = 1 - R(t2 - t1)$$

En considérant $N(0)$ le nombre d'entités fonctionnant à l'instant $t = 0$ et $N(t)$ le nombre d'entités survivant à l'instant t , le quotient

$$N(t)/N(0)$$

représente une mesure expérimentale de la fiabilité des entités identiques.

Exemple

On considère le nombre de défaillances observées sur les diodes laser utilisées pour la lecture de disques optiques numériques pour un parc de 270 lecteurs (tableau 1).

La figure représente la loi $R(t)$ approximée par une représentation avec 8 classes d'intervalles des temps.

Tableau 1 Statistiques des défaillances

Intervalle (h)	Nombre de défaillances par intervalle
0 à 500	50
500 à 1000	100
1000 à 1500	50
1500 à 2000	30
2000 à 2500	20
2500 à 3000	10
3000 à 3500	5
3500 à 4000	5
Total	270

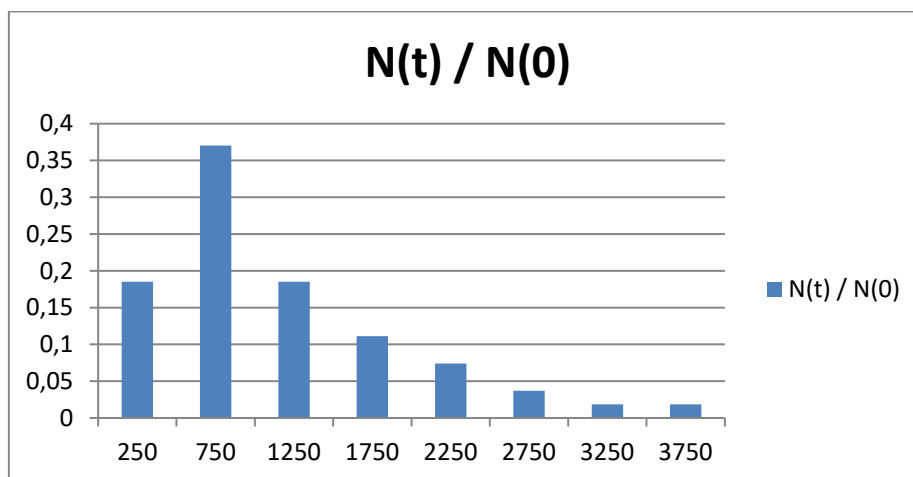


Figure 6 Taux de défaillance par intervalle de temps

1.3 Les mesures de la fiabilité

La "qualité" du matériel, du point de vue de la fiabilité, est donnée par un certain nombre d'indicateurs ou mesures de la performance.

1.4 Densité de défaillance

La densité de probabilité de l'instant de la défaillance T s'obtient en dérivant la fonction de répartition $F(t)$:

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

Ainsi, lors d'une approche expérimentale [4], $F(t)$ et $f(t)$ ont les valeurs suivantes :

$$F(t) = \frac{N - n(t)}{N}$$

$$f(t) = -\frac{1}{N} \cdot \frac{dn(t)}{dt}$$

Une autre mesure associée à la fiabilité c'est le temps moyen de fonctionnement jusqu'à la première défaillance, noté M.T.T.F. (" *Mean Time To Failure*"). Cette notion permet notamment de visualiser plus concrètement la fiabilité d'un matériel. Elle s'exprime par :

$$MTTF = \int_0^{\infty} t \cdot f(t) \cdot dt = - \int_0^{\infty} t \cdot \frac{dR(t)}{dt} \cdot dt$$

Le Mean Time To Failure est la durée moyenne de fonctionnement d'une entité avant la première défaillance et s'estime par:

$$\mu = \frac{T}{n}$$

1.5 Taux de défaillance :

Le taux de défaillance $\lambda(t)$ est défini comme la probabilité conditionnelle qu'une entité soit en panne à l'instant $t + \Delta t$ sachant qu'il était en fonction à l'instant t . Il arrive souvent que des entités présentent un taux de défaillance en fonction du temps suivant une courbe dite « en baignoire » (Figure 3).

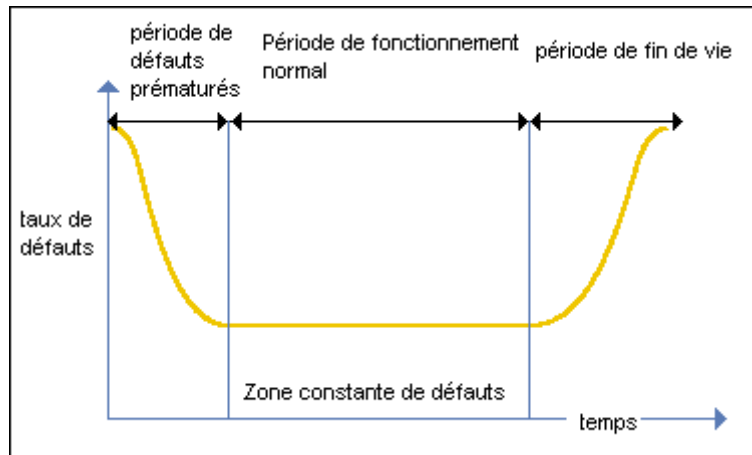


Figure 7 Courbe en baignoire

$$\lambda = \frac{n}{N * T}$$

Le taux de défaillance est applicable à tout composant non réparable. Il est associé au temps de la première défaillance et au temps moyen à la première défaillance.

Le modèle classique de « courbe en baignoire » caractérise la forme générale de l'évolution du taux de défaillance $\lambda(t)$ des composants depuis leur début de vie jusqu'à l'apparition des phénomènes d'usure.

Trois périodes sont observées sur la courbe de la Figure 3 :

- Période de défaillance précoce: période de rodage de l'entité pendant laquelle le taux de défaillance décroît rapidement. Ce phénomène résulte de l'élimination progressive des composants « faibles » de la population mise en service.
- Période de défaillance à taux constant: durée de vie utile de l'entité pendant laquelle les défaillances se produisent avec un taux passablement constant.
- Période de défaillance d'usure: période de dégradation de l'entité durant laquelle le taux de défaillance augmente rapidement. Elle est révélatrice des phénomènes de vieillissement (fatigue, usure, corrosion).

1.6 Application exemple de l'étude de la fiabilité

Dans une usine de textile, sur une machine de découpe de tissu, la rupture du support de la lame de découpe provoque des arrêts importants (changement du support, réinitialisation de la machine, réglages).

Le service maintenance décide d'étudier **la fiabilité de ces éléments** en vue d'instaurer une action de **maintenance préventive systématique** les concernant. Sachant que l'entreprise possède 9 machines de découpe.

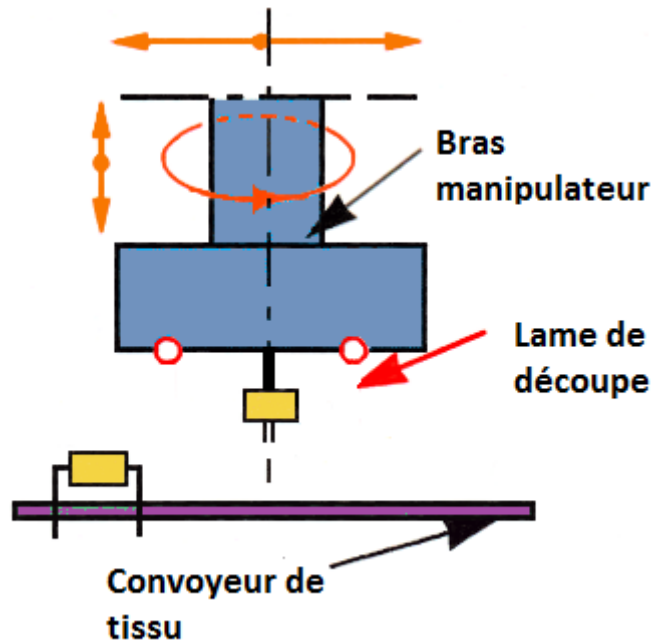


Figure 8 Schéma de la machine de découpe

Il faut d'abord commencer par une étude de la fiabilité, entre autre le Taux de Défaillance :

Nombre de matériels défaillants pendant un intervalle de temps Δt

$$\lambda(t) = \frac{\text{Nombre de matériels défaillants pendant un intervalle de temps } \Delta t}{\text{Nombre de matériels en service au début de } \Delta t \times t}$$

Tableau 2 Calcul du taux de défaillance exprimé en défaillance / heure

Intervalles Δt (en heures)	0 – 150	150 – 300	300 – 450	450 – 600	600 – 750
Nombre de matériels en service au début de Δt	9	8	6	5	3
Nombre de matériels défaillants pendant Δt	1	2	2	2	2
TAUX DE DEFAILLANCE (défaillance/heure) $\lambda(t)$	$7,4 \times 10^{-4}$	$16,7 \times 10^{-4}$	$22,2 \times 10^{-4}$	$33,3 \times 10^{-4}$	$66,6 \times 10^{-4}$

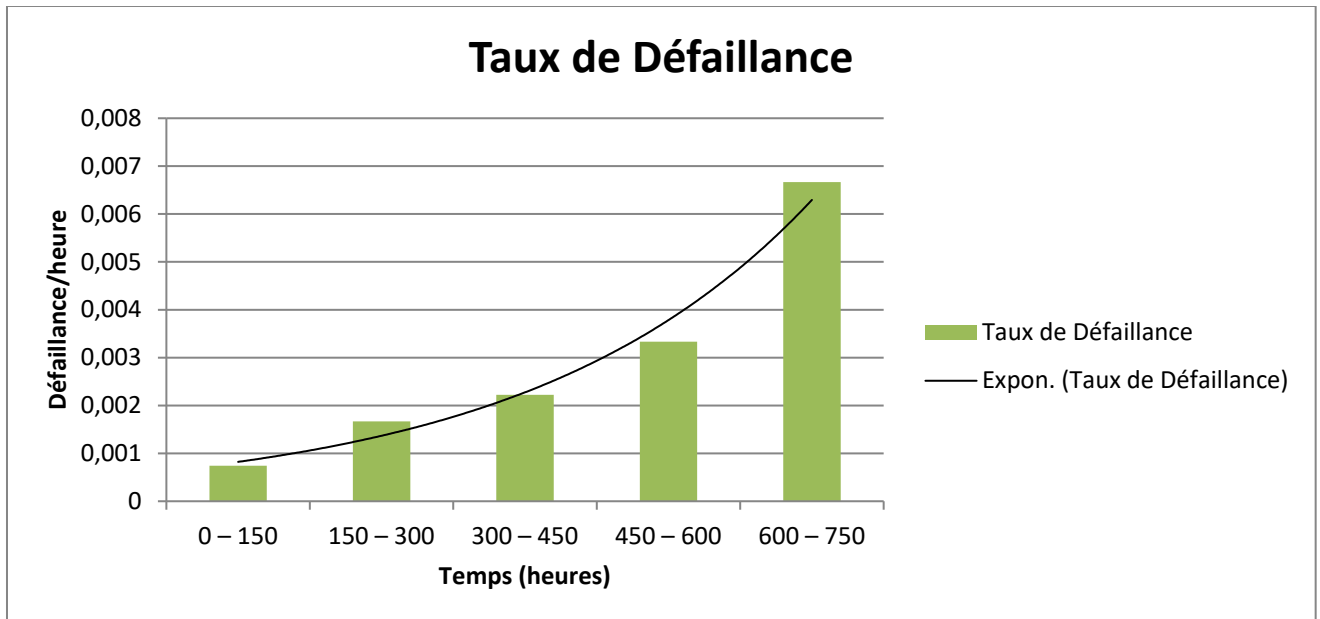


Figure 9 Représentation graphique du taux de défaillance en fonction du temps

Cette courbe met en évidence :

- La période de défaillance précoce
- La période de défaillance à taux constant
- La période de défaillance par vieillissement

Un changement systématique du support des lames est-il envisageable :

- Envisageable
- N'est pas envisageable

Une maintenance préventive systématique est alors envisageable.

Pour savoir quand les changer ces supports (la périodicité) il faut procéder à une étude de la Fiabilité.

La fonction fiabilité R(t) :

$$R(t) = \frac{\text{Nombre de matériels sans défaillances entre 0 et t}}{\text{Nombre total de matériels observé}}$$

Tableau 3 Calcul de la fonction fiabilité $R(t)$

Intervalles Δt (en heures)	0 – 150	150 – 300	300 – 450	450 – 600	600 – 750
Nombre de matériels défaillants dans l'intervalle Δt	1	2	2	2	2
Nombre de matériels sans défaillants à la fin de l'intervalle Δt	8	6	4	2	0
FIABILITE $R(t)$	0,89	0,75	0,67	0,5	0

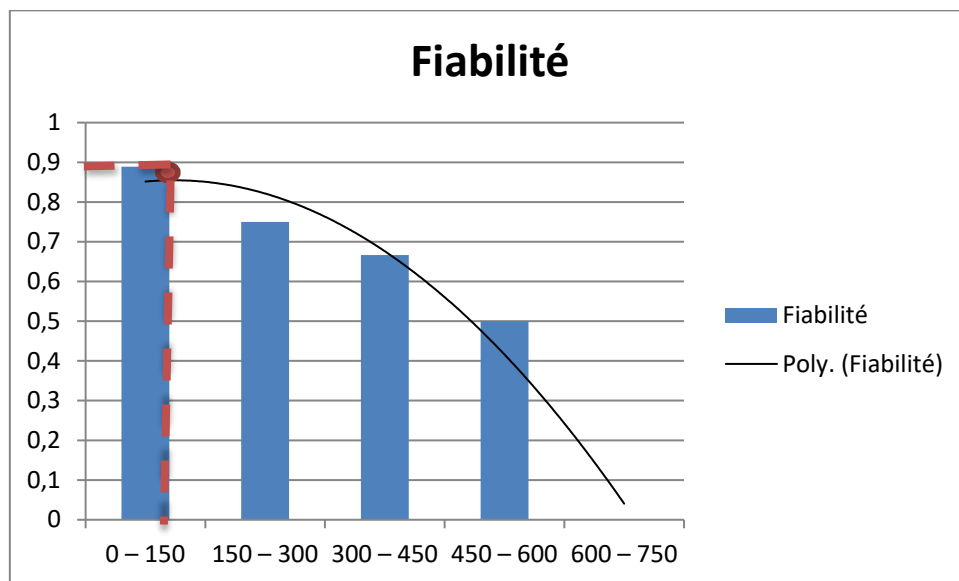


Figure 10 Représentation graphique du taux de Fiabilité en fonction du temps

Le taux de fiabilité acceptable pour un équipement de production qui entraîne l'arrêt total de l'entreprise ne doit pas être inférieur à 90%, de ce fait, le changement périodique des supports de la lame de découpe devra se faire toutes les 150 heures.

Etant donné que L'entreprise travaille 8 heures par jour, 5 jours par semaine (congé le Vendredi, Samedi et jours fériés) et ferme durant le mois d'août du 1 au 28.

La périodicité du changement est de :

$$T = 150 \text{ h} = 150/8 \text{ jours} = 18 \text{ jours}$$

Définition des dates de changement de support sur le premier trimestre de l'année 2017, sachant que le dernier changement a été effectué en 26 Décembre 2016.

الأول	
23	سبتمبر 2015
24	عبد الأقصى
الثاني	
14	نوفمبر 2015
	الشمسية
23	نوفمبر 2015
01	عبد النور
الثالث	
29	نوفمبر 2015
	الطلال
23	نوفمبر 2015
	الإختبارات 1
الرابع	
01	ديسمبر 2016
	الشمسية
28	ديسمبر 2016
	الطلال
الخامس	
01	يناير 2016
15	يناير 2016
	الإختبارات 3

نوفمبر 2015												
1	2	3	4	5	6	7	8	9	10	11	12	13
14	15	16	17	18	19	20	21	22	23	24	25	26
27	28	29	30									
ديسمبر 2015												
1	2	3	4	5	6	7	8	9	10	11	12	13
14	15	16	17	18	19	20	21	22	23	24	25	26
27	28	29	30	31								
يناير 2016												
1	2	3	4	5	6	7	8	9	10	11	12	13
14	15	16	17	18	19	20	21	22	23	24	25	26
27	28	29	30	31								
فبراير 2016												
1	2	3	4	5	6	7	8	9	10	11	12	13
14	15	16	17	18	19	20	21	22	23	24	25	26
27	28	29	30	31								
مارس 2016												
1	2	3	4	5	6	7	8	9	10	11	12	13
14	15	16	17	18	19	20	21	22	23	24	25	26
27	28	29	30	31								
أبريل 2016												
1	2	3	4	5	6	7	8	9	10	11	12	13
14	15	16	17	18	19	20	21	22	23	24	25	26
27	28	29	30	31								
ماي 2016												
1	2	3	4	5	6	7	8	9	10	11	12	13
14	15	16	17	18	19	20	21	22	23	24	25	26
27	28	29	30	31								
يونيو 2016												
1	2	3	4	5	6	7	8	9	10	11	12	13
14	15	16	17	18	19	20	21	22	23	24	25	26
27	28	29	30	31								



Figure 11 Dates du changement systématique des supports des lames de découpe.

2. La disponibilité

Se définit par la probabilité d'être en état d'accomplir sa fonction à un instant donné. La norme AFNOR X 60-500 définit la disponibilité comme « l'aptitude d'une entité à être en état d'accomplir une fonction requise dans des conditions données, à un instant donné ou pendant un intervalle de temps donné, en supposant que la fourniture des moyens extérieurs nécessaires de maintenance soit assurée ».

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

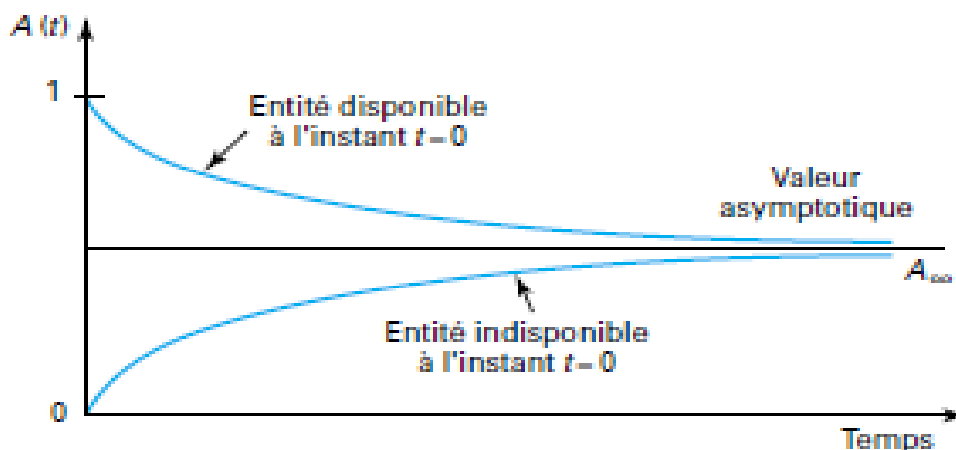


Figure 12 Evolution de la probabilité de disponibilité

2.1 Les grandeurs moyennes associées à la disponibilité :

Le temps moyen de disponibilité (TMD) ou durée de bon fonctionnement après réparation, ou mean up time (MUT) : durée moyenne de fonctionnement entre la réparation et la défaillance suivante ;

Le temps moyen d'indisponibilité (TMI) ou durée moyenne d'indisponibilité, ou mean down (MDT) : durée moyenne entre une défaillance et la remise en état suivante ;

La durée moyenne entre défaillance notée MTBF (mean time between failure) : durée moyenne entre deux défaillances consécutives de l'entité. En général, on a la relation :

$$MTBF = MUT + MDT$$

La disponibilité asymptotique se déduit du MUT et du MTBF par la relation :

$$A_{\infty} = MUT / MTBF$$

Pour les industries disposant d'équipes performantes de maintenance, la valeur de A est supérieure à 80 %.

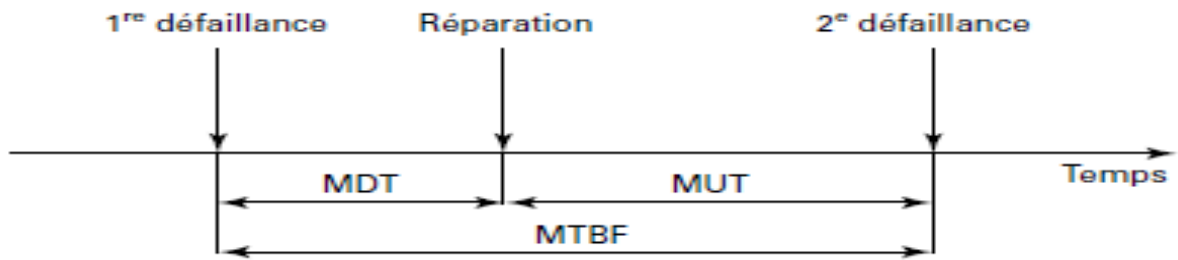


Figure 13 Les mesures de disponibilité

2.2 Disponibilité opérationnelle

L'indicateur de base de la disponibilité opérationnelle se calcule comme suit:

$$D_{op} = \frac{MTBF}{MTBF + MTTR} < 1$$

3. La maintenabilité

Selon la norme AFNOR c'est : « dans les conditions données d'utilisation, l'aptitude d'une entité à être maintenue ou rétablie, sur un intervalle de temps donné, dans un état dans lequel elle peut accomplir une fonction requise, lorsque la maintenance est accomplie dans des conditions données, avec des procédures et des moyens prescrits ».

Au sens mathématique, la maintenabilité se mesure par la probabilité que la maintenance d'une entité accomplie dans des conditions données, avec des procédures et des moyens prescrits, soit achevée au temps t, sachant que l'entité est défaillante au temps t=0.

$$M(t) = P [\text{la maintenance de E est achevée au temps t}]$$

3.1 Mesures liées à la maintenabilité

MTTR Le Mean Time To Repair est la durée moyenne de réparation et s'estime par:

$$MTTR = \frac{\sum_{i=1}^n MTTR_i}{n}$$

Où $MTTR_i$ représente le MTTR de la ième défaillance.

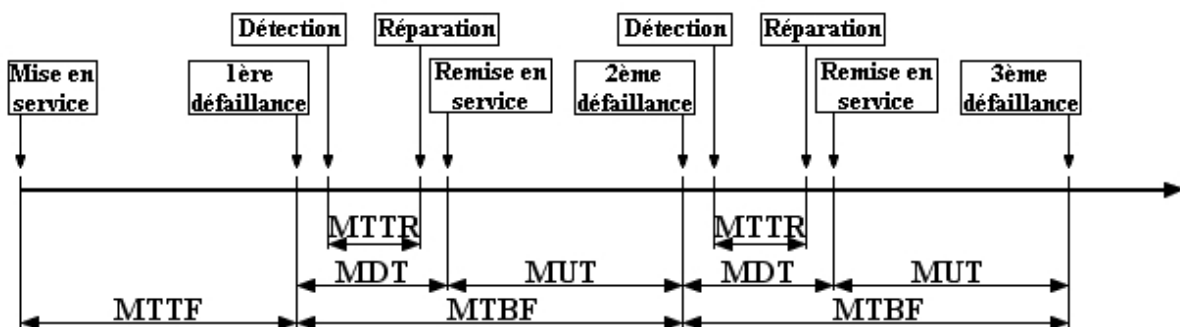


Figure 14 Les durées caractéristiques de Fiabilité-Maintenabilité-Disponibilité

Remarque : La maintenabilité est un équivalent à la fiabilité mais appliqué à la réparation au lieu de la défaillance.

3.2 Taux de réparation $\mu(t)$

la probabilité que l'entité soit réparée entre t et $t + dt$ sachant qu'elle n'était pas réparée sur l'intervalle $[0, t]$.

$$\mu(t) = \frac{1}{1 - M(t)} \frac{dM(t)}{dt}$$

4. Analyse de fiabilité, maintenabilité, disponibilité possibles

Tableau 4 Ce tableau indique les analyses correspondant au type de système [5].

Arborescence d'un système	Caractéristique	Analyses correspondantes		
		R(t)	M(t)	D(t)
Système	Réparable	<i>MTBF</i>	<i>MTTR</i>	D_{op}
	Non réparable	<i>MTTF</i>	X	X
Sous-système interchangeable	Réparable	<i>MTBF</i>	<i>MTTR</i>	X
	Consommable ¹	<i>MTTF</i>	X	X
Composant	Consommable	<i>MTTF</i>	X	x

5. La sécurité

Définition : Traduit l'aptitude d'une entité à éviter de faire apparaître, dans des conditions données, des événements critiques ou catastrophiques. Implique actuellement les aspects réglementaires de la sécurité des personnes.

Sûreté : La mesure dans laquelle un système fonctionne ou défaille sans incidence «catastrophique » sur son environnement.

Sécurité : La mesure dans laquelle un système résiste à des fautes intentionnelles.

Son évaluation découle de l'analyse des défaillances car elle concerne la probabilité de défaillance dangereuse, non atteignabilité d'état dangereux.

Une analyse prévisionnelle de sûreté de fonctionnement est un processus d'étude d'un système réel de façon à produire un modèle abstrait du système relatif à une caractéristique de sûreté de fonctionnement (fiabilité, disponibilité, maintenabilité, sécurité). Les éléments de ce modèle seront des événements susceptibles de se produire dans le système et son environnement, tels par exemple :

- Des défaillances et des pannes des composants du système,
- Des événements liés à l'environnement,
- Des erreurs humaines en phase d'exploitation.

Le modèle permet ainsi de représenter toutes les défaillances et les pannes des composants du système qui compromettent une des caractéristiques de SdF.

Afin d'aider l'analyste, plusieurs méthodes d'analyse ont été mises au point. Les principales sont :

APD Analyse Préliminaire des Dangers,

AMDE Analyse des Modes de Défaillances et de leurs Effets,

MDS Méthode du Diagramme de Succès,

MTV Méthode de la Table de Vérité,

MAC Méthode de l'Arbre des Causes,

MCPR Méthode des Combinaisons de Pannes Résumées,

MACQ Méthode de l'Arbre des Conséquences,

MDCC Méthode du Diagramme Causes-Conséquences,

Dans le chapitre suivant nous étudierons les modèles avancées de la sûreté de fonctionnement.

CHAPITRE 3 : Les modèles de SDF: Les modèles avancés

Les modèles avancés de la sûreté de fonctionnement se basent sur la probabilité élémentaire et surtout à variables aléatoires continues. Nous commencerons par un aperçu sur la matière.

La probabilité a différentes formes pour différentes interprétations :

La probabilité dite « **classique** » qui formalise les jeux de hasard et qui représente le rapport du nombre de cas favorables sur le nombre de cas possibles. Elle est donc fondée sur le dénombrement. Exemple :

Dans une classe de 32 étudiants avec 18 de sexe masculin et les reste de sexe féminin. La probabilité de tirer une fille est : $p(\text{fille}) = \frac{\text{Card}(\text{fille})}{\text{Card}(\text{etudiant})} = \frac{14}{32} = 0.43 = 43\%$

En étendant cette définition avec des cas non-équiprobables c'est-à-dire lorsqu'un cas est plus favorable qu'un autre, exemple d'un dé truqué (plus lourd d'un côté qu'un autre), c'est la probabilité dite « **parfréquence** ». Cette interprétation est fondée sur des « tirages » (ou « essais », « expériences ») que l'on peut répéter indéfiniment et à l'identique (par exemple, un lancer de dé ou le tirage d'une carte). La probabilité d'un évènement est alors définie par sa fréquence relative d'occurrence au cours d'une longue série (qui tend vers l'infini) de ces « tirages ».

Toujours avec l'exemple du lancement du dé, il y'a 1/6 de chance d'obtenir un AS, cependant sur 9000 lancements on n'aura pas 1500 AS mais un nombre qui s'y approche. De ce fait la probabilité fréquentiste est la limite entre les statistiques et la probabilité :

$$p(AS) = \frac{\text{Nombre de fois d'obtention d'un AS}}{\text{Nombre de lancement}}$$

NB : Les liens entre probabilité et statistique sont forts. Les statistiques partent de la réalité d'une population et cherchent à la modéliser avec des lois mathématiques pour l'expliquer et/ou extrapoler son comportement à une autre population. Les probabilités définissent les lois mathématiques auxquelles obéissent des expériences régies par le hasard.

Enfin, une probabilité « **subjective** » telle qu'on parle d'une estimation de la possibilité, on retrouve souvent le terme de « degrés de croyance » (ou « l'état des connaissances »). Cette interprétation est également appelée « bayésienne », du nom de Thomas Bayes qui développa, au XVIIIème siècle, le concept de probabilité a posteriori, permettant d'actualiser une probabilité a priori sur la base de nouvelles informations.

Très utilisée en Ingénierie autant qu'outil d'aide à la décision et qui doit être actualisée en fonction du retour d'expérience ; tandis que l'interprétation fréquentiste des probabilités est

plus idéaliste, plus favorable au mathématicien (statisticien) pour qui le traitement des (nombreuses) données disponibles permet d'en déduire des tendances sur lesquels miser (les assurances, la finance).

Dans un contexte d'étude de sûreté, les occurrences d'évènements tels que des défaillances de systèmes ou des erreurs humaines sont souvent le fait de caractéristiques ou de conditions environnementales qui leurs sont propres, rendant alors le concept de « tirages identiques », et donc l'interprétation fréquentiste, inappropriés. De plus, les évènements à traiter étant généralement (et heureusement) rares, et même parfois jamais observés, les probabilités manipulées sont plus justement interprétées comme étant subjectives. En particulier, le manque de données appropriées et le recours (indispensable) à des « jugements » implique cette subjectivité [6, 7].

1. Application de l'approche classique de probabilité dans la sûreté de fonctionnement

La fiabilité représente l'aptitude d'une entité à accomplir une fonction requise E, dans des conditions données, pendant une durée donnée. Elle est généralement mesurée par la probabilité qu'une entité accomplisse la fonction requise E, pendant un intervalle de temps $[0, t]$. La théorie de fiabilité repose sur deux hypothèses fondamentales :

- Hypothèse probabiliste : le fonctionnement du système est complètement caractérisé par des mesures de probabilité.
- Hypothèse d'états binaires : le système ne peut avoir que deux états de fonctionnement; l'état de défaillance et l'état de fonctionnement normal.

Ce formalisme correspond à la fiabilité PROBIST [8].

Lors du calcul prévisionnel de la fiabilité des systèmes, on utilise souvent des bases de données connues, mais il est préférable, quand cela est possible de recueillir les données de retours d'expériences auprès des fabricants des composants que l'on utilise. Cependant, ces données sont difficiles à obtenir pour des composants mécaniques. Les constructeurs ne s'efforcent pas de les collecter systématiquement, du fait qu'il est difficile de trouver une métrique de sûreté de fonctionnement. Dans le cas où elles existeraient, elles sont conservées confidentiellement.

Les recueils de données de fiabilité les plus connus pour des dispositifs électroniques et mécaniques sont présentés dans les tableaux suivants :

Dans la majorité des recueils, les données de fiabilité sont fournies sous forme de taux de défaillance constants principalement pour les composants électroniques et sous forme de

durées de vie moyennes, valeurs minimales et maximales ou de probabilité de bon fonctionnement, en particulier pour les composants mécaniques.

Ces recueils sont mis à jour régulièrement pour tenir compte des évolutions technologiques.

Tableau 5 **Recueils de données de fiabilité en électronique [9]**

<i>Source</i>	<i>Titre</i>	<i>Editeur</i>	<i>Dernière version</i>
IEEE STD	IEEE Guide to the Collection and Presentation of Electrical, Electronic Sensing Component and Mechanical Equipment Reliability Data for Nuclear Power Generating Stations.	Institution of Electrical and Electronic Engineers, New York, USA.	IEEE STD500,1984.
MIL-HDBK-217	Military Handbook-Reliability Prediction of Electronic Equipment.	United States Department Of Defense.	MIL-HDBK-217F, notice 2, 28 Février 1995.
EIREDA	European Reliability Industry Data Handbook	European Commission and Electricité de France CRETE UNIVERSITY PRESS	1998
EPRD	Electronic Parts Reliability Data.	Reliability Analysis Center, RAC,NEW YORK, USA.	EPRD 97, 1997.
GJB	ChineseMilitary Standard.	Beijing Yuntong Forever Sci.-Tech.	GJB/Z229B,1998
RDF (CNET)	Recueil de Données de Fiabilité.	Centre National d'Etudes des Télécommunications, UTE, Paris, France.	RDF 2000-UTE. C80-810, Juillet 2000.
Telcordia/Bellcore	Reliability Prediction Procedure for Electronic Equipment.	Telcordia Technologies, New Jersey, USA.	Telcordia SR-332 Issue 1, Mai 2001.
OREDA	Onshore and offshore Reliability Data	OREDA is a project organization sponsored by eight oil and gas companies with worldwide operations.	2000

Exemple : La probabilité pour qu'une vanne de déluge (système d'incendie) ne s'ouvre pas est de : $5.22 * 10^{-3}$ (ref EIREDA).

La détermination de la fiabilité d'un système électronique, mécanique ou autre nécessite tout d'abord de connaître la loi de la fiabilité (ou la loi de défaillance) de chacun des composants intervenant dans le système. Ceci est simple pour certains types de systèmes tels que les systèmes électroniques, or ce n'est pas le cas pour des systèmes mécaniques à cause de la complexité de la structure du système étudié.

Pour un système électronique chaque composant a un poids important dans la fiabilité du système, la fiabilité du système est donc calculée en fonction de la fiabilité de tous ses composants.

La détermination des taux de défaillance des composants est effectuée soit à partir des modèles développés dans des bases de données disponibles, soit à partir d'essais effectués sur les composants ou bien à partir des résultats d'exploitation des produits.

La fiabilité d'un système mécanique, contrairement à l'électronique, repose sur la fiabilité de quelques composants élémentaires responsables de son dysfonctionnement, dits composants « responsables » ou « critiques » (parfois un seul), contribuant presque totalement à la probabilité de défaillance de l'ensemble. Les autres composants pouvant être considérés de probabilité de défaillance pratiquement nulle.

L'identification de ces composants se fait en effectuant des analyses qualitatives, telles que l'APR (Analyse Préliminaire des Risques) et l'AMDEC (Analyse de Modes de défaillance, de leurs Effets et de leurs Criticité), et des analyses quantitatives telles que l'analyse par arbres de défaillance (voir Annexe A). Ensuite, nous en créerons un modèle à partir d'un diagramme de fiabilité, par exemple sous forme d'un schéma bloc. Le taux de défaillance du système sera on calculé en fonction de l'architecture du système donnée par ce schéma.

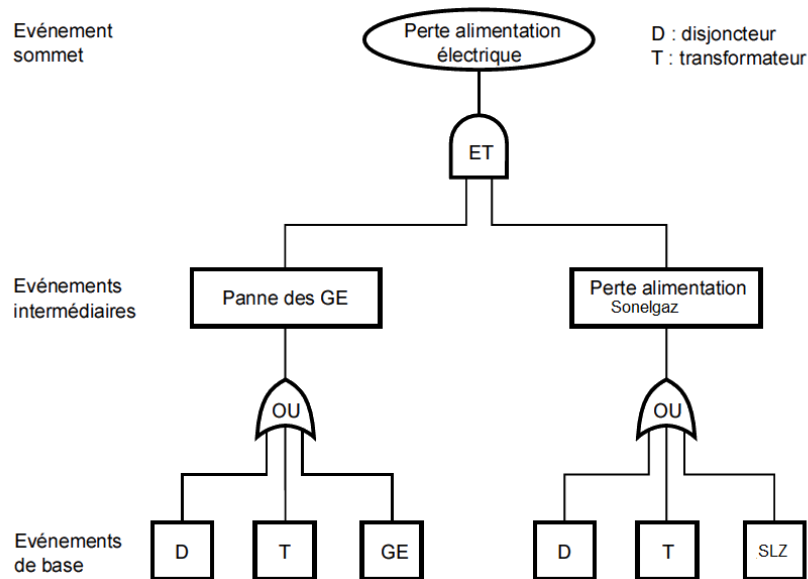


Figure 15 Exemple d'un arbre de défaillance de l'alimentation électrique

Le plus souvent, les systèmes mécaniques sont considérés à configuration série, Du fait qu'on gardé que les composants critiques. Dans ce cas et lorsque les composants du système sont supposés indépendants, les caractéristiques de fiabilité sont les suivantes :

- La fiabilité $R(t)$ d'un système mécanique non réparable est le produit des fiabilités des composants "critiques", la défaillance de l'un de ces éléments entraîne la défaillance du système:

$$R(t) = \prod_{i=1}^n R_i(t)$$

- Le taux de défaillance $\lambda(t)$ du système est approximativement égal à la somme destaux de défaillance individuels ($\lambda_i(t), i=1, \dots, n$) :

$$\lambda(t) = \sum_{i=1}^n \lambda_i(t)$$

Il y'à aussi les essais de fiabilité. Les essais de fiabilité sont destinés à vérifier ou compléter les données de fiabilité existantes ou à les élaborer lorsqu'elles ne sont pas disponibles, et ceci en procédant à des tests sur un certain nombre d'entités.

La norme UTE C20-321 [Afnor, 1988] définit les différents types d'essais de fiabilité :

Les essais de fiabilité sont en général des essais de conformité en fiabilité. Ils s'appliquent en général dans l'un des cas suivants :

- Modèles d'étude ou prototypes,
- Lot de présérie,

- Lot (ou lots) de production.

Quant à leur réalisation, différentes mesures sont prises :

- La population doit être homogène : les dispositifs sont de même type, produit par les mêmes moyens et dans des conditions similaires, afin que les essais soient représentatifs,
- Le choix des dispositifs à mettre en essai doit être fait dans la population concernée selon un échantillonnage aléatoire,
- Durant le déroulement des essais les dispositifs sont exposés aux mêmes contraintes que les équipements livrables.

Des essais de différents types peuvent être effectués. Nous les distinguons en fonction:

- De leur objectif qui peut être soit la mesure de taux de défaillance et l'ajustement d'un modèle de loi de fiabilité, ou le contrôle d'homologation, de maintenance et réduction du temps de fabrication, ou bien faire des analyses technologiques des défaillances et maîtrise des causes de défauts.
- Des contraintes appliquées qui peuvent être constantes (normales ou fortes), cycliques ou échelonnées.
- De la procédure ou des résultats attendus. Sur un échantillon de dispositifs de taille définie, on choisit à l'avance : un critère d'arrêt (l'essai est terminé lorsque tous les dispositifs sont défaillants : essai complet), une durée limitée (l'essai est arrêté au bout d'un temps fixé à l'avance, avec ou sans remplacement des produits défectueux : essai tronqué ou censure de type I), un nombre de défaillance limité (l'essai est arrêté au bout d'un nombre de défaillances fixé à l'avance, avec ou remplacement des produits défectueux : essais censurés de type II), ou pas de critère préalable relatif à la durée de l'essai ou au nombre maximal de défaillances (l'arrêt n'est décidé qu'en fonction des résultats cumulés au cours de l'essai : essai progressif ou séquentiel).
- Du type de produit. Composant ou équipement.
- Du coût d'essai. Le coût d'une campagne d'essais dépend de deux paramètres essentiels : le nombre d'essai et le temps d'un essai. L'objectif industriel étant de faire des essais aux moindres coûts, il faudrait donc soit réduire les tailles d'échantillons ou réduire les durées d'essais; pour chaque solution est accordé un certain type d'essai.

2. Application de l'approche Fréquentiste dans la sureté de fonctionnement

Lorsqu'on s'intéresse à la tendance d'un facteur du système, fiabilité, disponibilité..., sur une durée passée, actuelle ou future, on s'intéresse à la distribution d'une variable exp : taux de défaillance. L'approche fréquentiste est alors une extension de l'approche classique.

Ces facteurs seront alors représentés par des fonctions à répartition d'une variable aléatoire continue.

2.1 Définition de la fonction de la fiabilité en loi probabilistes

La fiabilité est la probabilité que l'entité considérée ne tombe pas en panne avant un instant t donné :

$R(t) = P(E \text{ non défaillante sur la durée } [0, t], \text{ en supposant qu'elle n'est pas défaillante à l'instant } t = 0).$

$R(t) = P(T > t)$ (comme vu dans le chap 2 .1. 1. 1)

Allure de $R(t)$ loi exponentielle $R(t) = \exp(-\lambda t) \lambda > 0$

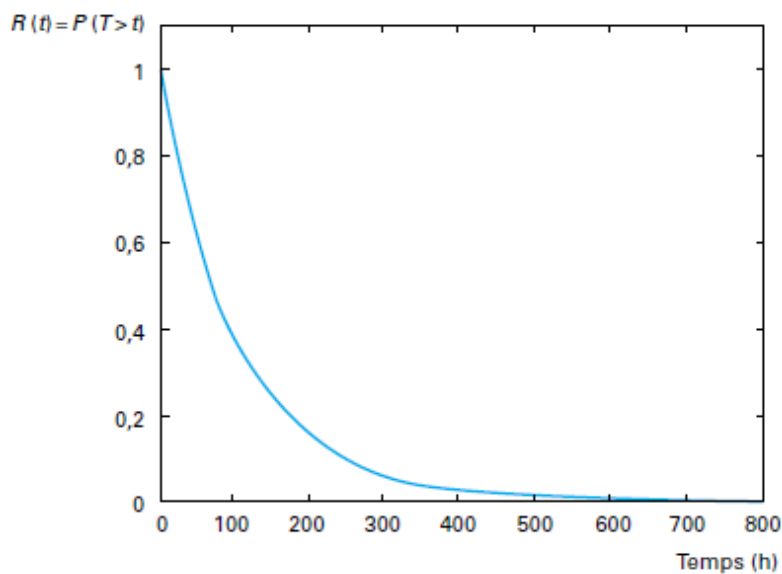


Figure 16 Allure de $R(t)$ loi exponentielle

avec λ paramètre de la loi exponentielle.

La valeur moyenne et la variance sont données par :

$$m = \frac{1}{\lambda}$$

Et

$$var(x) = \frac{1}{\lambda^2}$$

2.2 Relations entre $R(t)$, $F(t)$, $f(t)$ et $\lambda(t)$

L'application des propriétés de la fonction logarithme népérien et du calcul intégral conduit aux relations suivantes :

$$d(\ln(u)) = \frac{du}{u} = \int \frac{du}{u} = \ln(u) \Rightarrow u = \exp \int \frac{du}{u}$$

et

$$\lambda(t) = -\frac{R'(t)}{R(t)} = -\frac{d[\ln(R(t))]}{dt}$$

$$\int_0^t \lambda(u) du = -[\ln R(u)]_{t_0} = -\ln(R(t))$$

Car $R(0)=1 \Rightarrow \ln(R(0))=0 \Rightarrow R(t) = \exp - \int_0^t \lambda(u) du$

Comme on mesure la fiabilité à partir de $t = 0$, il est évident que $R(0) = 1$ (pas de défaillance à $t = 0$), d'où les relations entre $F(t)$, $R(t)$, $f(t)$ et $l(t)$ définies dans le tableau suivant.

Tableau 6 relations entre $F(t)$, $R(t)$, $f(t)$ et $l(t)$

	F(t)	R(t)	f(t)	$\lambda(t)$
F(t)		$1-R(t)$	$\int_0^t f(u) du$	$1 - \exp - \int_0^t \lambda(u) du$
R(t)	$1-F(t)$		$\int_t^\infty f(u) du$	$\exp - \int_0^t \lambda(u) du$
f(t)	$\frac{dF(t)}{dt}$	$-\frac{dR(t)}{dt}$		$\lambda(t) \exp \int_0^t \lambda(u) du$
$\lambda(t)$	$\frac{\frac{dF(t)}{dt}}{1 - F(t)}$	$-\frac{R'(t)}{R(t)}$	$\frac{f(t)}{\int_t^\infty f(u) du}$	

L'application de ces relations aux lois exponentielle et de Weibull conduit aux expressions des taux de défaillances dans le tableau suivant.

Tableau 7 lois exponentielle et de Weibull

Loi	Exponentielle	Weibull
f(t)	$\lambda \exp - \lambda t$	$\beta \frac{(t - \gamma)^{\beta-1}}{n^\beta} \exp - \left(\frac{t - \gamma}{n}\right)^\beta$
$\lambda(t)$	λ	$\beta \frac{(t - \gamma)^{\beta-1}}{n^\beta}$

L'étude de la loi de Weibull, qui dépend de trois paramètres, est très instructive pour représenter des taux de défaillance décroissants, constants ou croissants en ne faisant varier

que le paramètre β . Comme le représente la figure 9, le taux de défaillance est constant pour $\beta = 1$, croissant pour $\beta > 1$ et décroissant si $\beta < 1$. Si le paramètre β varie au cours du temps sous la forme d'une expression linéaire croissante avec le temps et à partir d'une valeur initiale inférieure à 1, on observe une décroissance du taux de défaillance, suivie d'un plateau et d'une remontée. Cette allure est typique du comportement des systèmes qui s'usent au cours du temps.

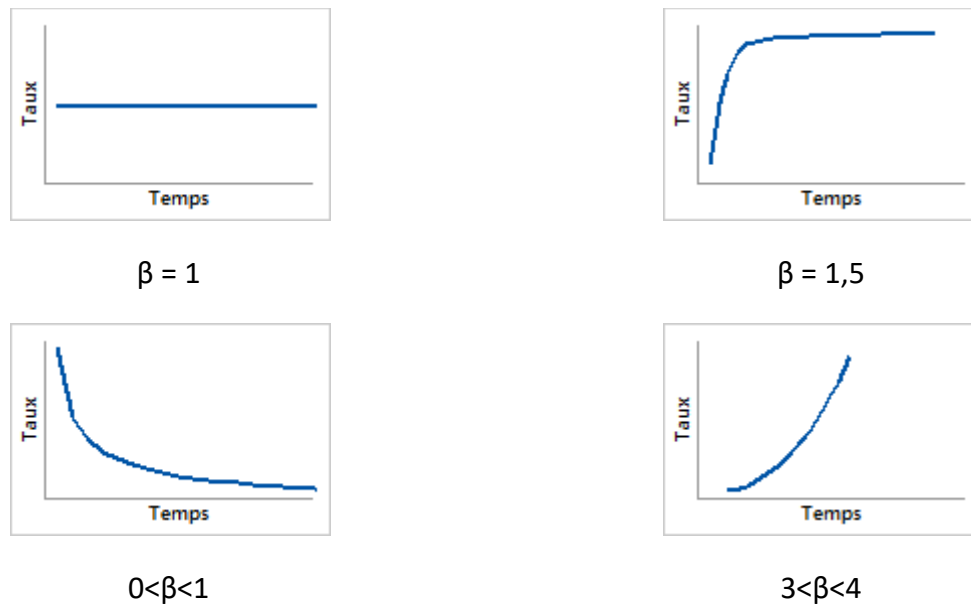


Figure 17 Allure de $R(t)$ loi Weibull

Ce qui permet de retrouver la courbe en baignoire vu dans le chapitre 2

3. Application de l'approche subjective dans la sûreté de fonctionnement

Lorsqu'on souhaite faire une étude de fiabilité prévisionnelle et lorsque les données sont insuffisantes et incertaine, l'approche bayésienne apporte une réponse.

Comme pour les essais vus dans la section précédente, les techniques bayésiennes sont utilisées pour réduire la taille des essais d'estimation, améliorer l'estimation des paramètres de la fiabilité du produit par l'intégration du passé (données disponibles sur le produit concerné) et traiter le cas « zéro » défaillance observée, difficile à traiter avec une approche statistique classique qui demanderait une taille d'échantillon grande.

Il existe de nombreux travaux traitant les techniques bayésiennes. Nous pouvons citer les travaux fondamentaux : [10], [11],...et les travaux concernant les applications à la fiabilité : [12], [13].... .

Dans un plan d'essai bayésien, les résultats d'essais réalisés au cours du cycle de développement sont combinés avec un modèle de fiabilité « a priori », pour obtenir un modèle « a posteriori », par le biais du théorème de Bayes. Le modèle a priori est construit à l'aide des informations connues, acquises avant la mise en œuvre du premier essai (avis d'expert, informations provenant des études préalables telles que les analyses prévisionnelles, les AMDEC,...).

Les essais de fiabilité classiques sont destinés à fournir des échantillons nécessaires pour estimer des caractéristiques de fiabilité (les paramètres d'une loi de vie, taux de défaillance, etc.). Les informations contenues dans ces échantillons sont appelées informations objectives.

En plus des informations objectives, on dispose parfois d'informations supplémentaires de nature subjectives (informations subjectives) telles que des données sur des matériels de technologies similaires à celle du matériel que l'on étudie.

S'il est possible de traduire cette information subjective en attribuant aux paramètres à évaluer, une distribution de probabilité « a priori », il sera possible, en utilisant le théorème de Bayes, de combiner l'information subjective et l'information objective afin de construire une distribution « a posteriori » des paramètres considérés. Il est indispensable d'utiliser un modèle « a priori » simple permettant un traitement analytique du modèle « a posteriori ».

Dans le chapitre suivant, nous montrerons les fondements de l'approche Bayésienne et leurs applications.

CHAPITRE 4 : Les bases de l'approche bayésienne

1. Introduction

Thomas Bayes Né environ en 1702 à Londres - mort le 7 avril 1761 à Tun bridge Wells, dans le Kent) est un mathématicien britannique et pasteur de l'Église presbytérienne, connu pour avoir formulé le théorème de Bayes.

Ses découvertes en probabilités ont été résumées dans son Essai sur la manière de résoudre un problème dans la doctrine des risques (Essay Towards Solving a Problem in the Doctrine of Chances - 1763) publié - par un de ses amis, Richard Price - à titre posthume dans les comptes-rendus de l'Académie royale de Londres (The Philosophical Transactions of the Royal Society).

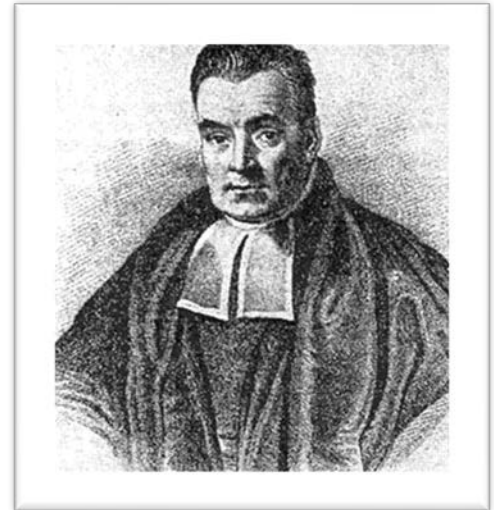


Figure 18 Thomas Bayes

2. La probabilité conditionnelle

On appelle probabilité conditionnelle la probabilité qu'un événement B se produise sachant que l'événement A s'est déjà produit. On la note $P(B|A)$ ou $P_A(B)$ et on la lit «probabilité que B se réalise sachant que A s'est produit». On l'appelle aussi probabilité à postériori de B sachant A.

Les probabilités conditionnelles prennent en compte les informations concernant l'issue d'une expérience qui modifie la probabilité des événements liés à cette expérience.

$P(B|A) = \frac{P(A \cap B)}{P(A)}$ Mathématiquement, soient (Ω, E, P) un espace probabilisé ou Ω est l'univers, E l'ensemble d'événements de probabilité non nulle et P la mesure. À tout événement B de A, nous associons le nombre noté $P(B|A)$ défini par:

Soit $B \in A$, $P(B|A)$ doit être proportionnel à $P(B)$ et $P(A|A)=1$

Soit $B \in A_c$ alors $P(B|A)=0$

Ceci nous mène à

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

Nous pourrions vérifier que l'application

$P(A)$ définie par $B \rightarrow p(B|A)$

Puis $P(A|B)$ appelé fonction de vraisemblance de A ou probabilité à priori de A sachant B.

2.1 Espérance conditionnelle

Soit (Ω, \mathcal{B}, P) un espace probabilisé, X une variable aléatoire réelle intégrable et B un évènement. On appelle espérance conditionnelle:

$$E(X|B) = \frac{1}{P(B)} \int_B X dp$$

$E(X|B)$ est l'espérance de X sachant que B s'est réalisé.

2.2 Densité conditionnelle

Soit (Ω, \mathcal{B}, P) , et soient X et Y deux variables aléatoires définies sur cet espace. Si l'on suppose que leur loi jointe peut être définie par une densité bi variable $f(x, y)$, et si de plus un y_0 vérifie $\int f(x, y_0) dx \neq 0$, alors il existe une loi absolument continue dont la densité est donnée par l'expression

$$g(x|y_0) = \frac{f(x, y_0)}{\int f(x, y_0) dx}$$

Cette fonction $g(x|y_0)$ est appelée : densité conditionnelle de X sachant $Y = y_0$. Intuitivement, cette expression peut être interprétée comme une formulation continue du théorème de Bayes.

2.3 Représentation par un arbre pondéré

Soient deux évènements A et B. On peut représenter par un arbre pondéré les probabilités suivantes lorsque l'on connaît les probabilités de B ou B lorsque A est réalisé.

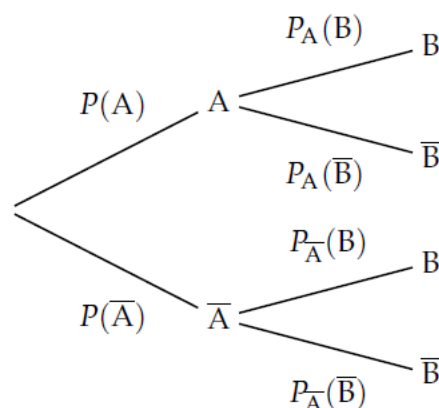


Figure 19 Arbre pondérée

Pour remplir et utiliser un arbre, on a les propriétés suivantes :

- Sur chaque branche de l'arbre, on écrit les probabilités correspondantes (attention pas de pourcentage).
- La somme des probabilités inscrites sur les branches issues d'un même nœud est égale à 1 (loi des nœuds).
- Le produit des probabilités inscrites sur chaque branche d'un chemin donne la probabilité de l'intersection des événements placés sur ce chemin.
- La probabilité d'un événement est la somme des probabilités des chemins qui aboutissent à cet événement

Exemple général:

Si je jette un dé non truqué, la probabilité d'obtenir un 6 est de $\frac{1}{6}$.

Si je lance ce même dé, qu'une tierce personne me cache le résultat et me dit « j'ai vu que le résultat est pair », la probabilité de l'événement « avoir un 6 » change. Je sais que les issues possibles se réduisent maintenant aux nombres 2, 4, 6. La probabilité d'obtenir un 6 devient donc $\frac{1}{3}$.

On dit que la probabilité d'obtenir 6, sachant que le nombre obtenu est pair, est de $\frac{1}{3}$.

Si A et B sont indépendants alors:

$$P(B|A) = p(B)$$

D'autre part, on peut noter que le théorème de Bayes permet d'écrire ceci:

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

Ce qui, en cas d'indépendance des événements A et B, se vérifie par:

$$P(B) = P(A) * \frac{P(B)}{P(A)}, \text{ ce qui est vrai.}$$

Exemple :

Supposons que nous ayons deux dés.

Lancer le premier dé, Quelle est la probabilité en lançant le deuxième dé pour avoir une somme minimal de 10 ?

La probabilité d'obtenir cette valeur minimale sachant la valeur du premier dé et totalement différente de la probabilité de l'avoir en lançant les deux dés à la fois:

$A = \{\text{Le résultat du premier lancer est } \dots\}$

$B = \{\text{La somme est supérieur ou égale à } 10\}$

$A = \{5\} \rightarrow P(A) = 1/6$

$B = \{(4,6), (6,4), (5,5), (5,6), (6,5), (6,6)\}$

$$P(B) = \frac{1}{6+6} + \frac{1}{6+6} + \frac{1}{6+6} + \frac{1}{6+6} + \frac{1}{6+6} + \frac{1}{6+6} = \frac{1}{6}$$

$$B \cap A = \{(5,5), (5,6)\} \rightarrow P(B \cap A) = \frac{1}{6+6} + \frac{1}{6+6} = \frac{1}{18}$$

$$B|A = \{(5), (6)\} \rightarrow P(B|A) = \frac{1}{6} + \frac{1}{6} = \frac{1}{3}$$

Et en appliquant la loi de Bayés :

$$P(B|A) = \frac{P(A \cap B)}{P(A)} = \frac{\frac{1}{18}}{\frac{1}{6}} = \frac{1}{3} \text{ Ce qui est vérifié.}$$

Exemple de l'arbre pondéré :

Dans un atelier, il y a 2% de pièces défectueuses. On effectue un test pour savoir si on doit accepter ou refuser une pièce. On a observé que :

- Si la pièce est bonne, elle est acceptée par ce test à 96%.
- Si la pièce est défectueuse, elle est refusée par ce test à 97%.

Quel est le pourcentage de retour client ?

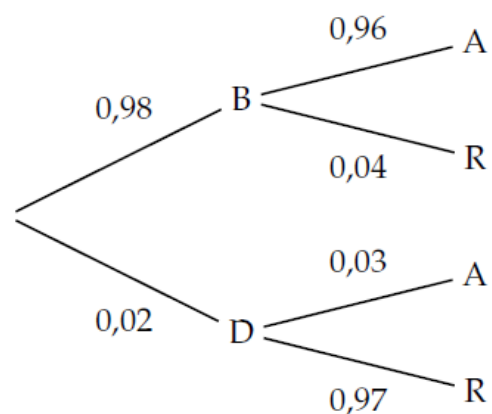
On appelle les événements suivants :

B : "la pièce est bonne",

D : "la pièce est défectueuse",

A "la pièce est acceptée",

R "la pièce est refusée".



On construit l'arbre ci-contre

On doit calculer la probabilité qu'une pièce soit défectueuse et acceptée :

$$P(D \cap A) = 0,02 \times 0,03 = 0,0006$$

On peut s'attendre à 0,06% de retour client.

3. Généralisation du théorème de Bayes

$P(A|B_i)$ est la probabilité conditionnelle d'obtenir l'événement A quand on sait que B_i en est la cause, $P(B_i)$ "Probabilité à priori de B_i , $P(A|B_i)$ la fonction de vraisemblance et $P(B_i|A)$ "Probabilité à postériori de B_i "

$$P(A) = P((B_1 \cup B_2 \cup \dots \cup B_i) \cap A) =$$

$$P((B_1 \cap A) \cup (B_2 \cap A) \cup \dots (B_i \cap A)) =$$

$$P(B_1 \cap A) + P(B_2 \cap A) + \dots + P(B_i \cap A) =$$

$$P(A|B_1)P(B_1) + P(A|B_2)P(B_2) + \dots + P(A|B_i)P(B_i) =$$

$$\sum_{i=1}^n P(A|B_i)P(B_i)$$

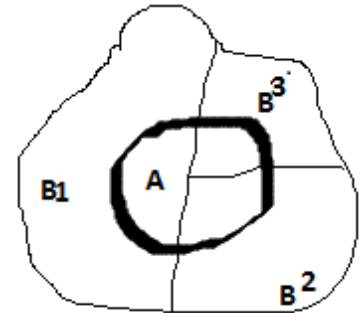


Figure 20 Imbrication des événements

On obtient ainsi la forme discrète du théorème de Bayes, d'où l'on déduit que, si l'on connaît l'ensemble des probabilités a priori pour que des événements causes B_i se produisent, ainsi que l'ensemble des probabilités que l'événement conséquence A se réalise en fonction des différents causes, on peut en déduire la probabilité qu'une cause particulière B_i soit à l'origine de l'occurrence de A constaté.

$$P(B_i|A) = \frac{P(A|B_i)P(B_i)}{\sum_{i=1}^n P(A|B_i)P(B_i)}$$

Exemple d'application : origine d'un matériel défaillant :

Trois sites fabriquent un même type des appareils électroménagers d'apparences identiques. Ils assurent respectivement 50%, 35%, 15% de la production.

On définit :

- l'événement C_i : une machine est produite sur le site S_i .
- l'événement E : une machine est défaillante sur une période garantie de 2 ans.

$P(C_i)$ représente la probabilité pour qu'un appareil d'être soit produit sur le site S_i .

L'ensemble des probabilités $P(C_i)$ représente la connaissance a priori sur l'origine de fabrication des matériels avant de connaître un éventuel retour d'expérience.

Les probabilités a priori sont respectivement :

$$P(C_1) = 0,5 ; P(C_2) = 0,35 ; P(C_3) = 0,15.$$

Des retours de garantie à 2 ans pour chaque site de fabrication ont donné les résultats

Suivants :

$$P(E|C_1) = 0,003 ; P(E|C_2) = 0,008 ; P(E|C_3) = 0,01.$$

$P(E|C_i)$ représentant la probabilité d'occurrence d'une panne quand on connaît la provenance de l'appareil.

Question : on souhaite connaître la probabilité pour un appareil non identifié d'avoir été fabriqué sur un site S_i , sachant qu'il a été constaté défectueux sur la période de garantie considérée.

En appliquant le théorème de Bayes sous sa forme discrète, on obtient respectivement :

$$P(C_1|E) = 0,26 ; P(C_2|E) = 0,48 ; P(C_3|E) = 0,26 .$$

Avec :

$$P(E) = 0,5 \cdot 0,003 + 0,35 \cdot 0,008 + 0,15 \cdot 0,01 = 0,00558.$$

On voit ici que la contribution la plus importante à l'événement indésirable E, panne d'une machine quelconque sur une période de garantie de 2 ans, est celle qui est due à la fabrication sur le site S_2 , bien que ce site ne fournisse pas la plus grande quantité de d'équipements machines, comme on peut le constater sur l'image suivante :

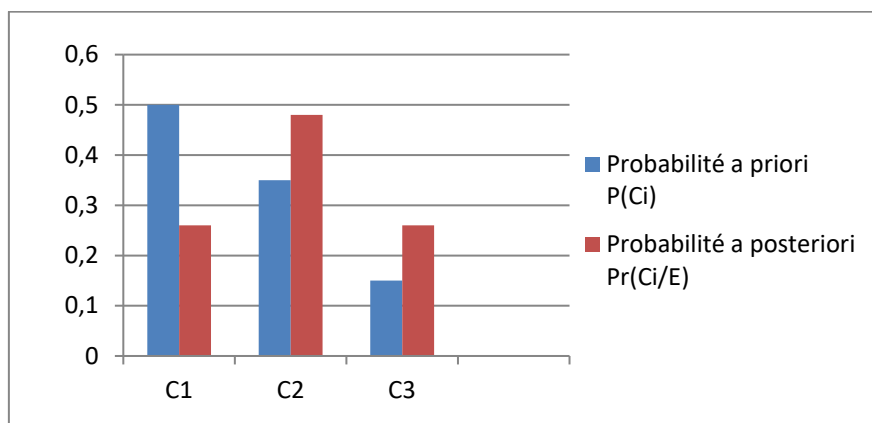


Figure 21 Probabilité a priori et a posteriori

Ainsi, la connaissance a priori a été modifiée par la connaissance de l'événement E conditionner par l'origine du site de fabrication ($E|C_i$).

Sans cette connaissance, on aurait pu penser intuitivement que, puisque le pourcentage de fabrication le plus élevé était $P(C_i)$, la machine défaillante provenait en priorité du site S_1 , puis de S_2 , puis de S_3 .

Lorsqu'on remarque l'apparition de l'événement E m fois tel que ($m > 0$), la formule de Bayes aura la forme suivante :

$$P(C_i | mE) = \frac{P(E|C_i)P(C_i)^m}{\sum_{i=1}^n P(E|C_i)P(C_i)^m}$$

4. Formulation continu du théorème de Bayes

L'occurrence d'un événement E (résultats d'expériences relatives à des essais, par exemple) dépend d'une variable aléatoire Θ (taux de défaillance λ , par exemple) qui lui est associée. On découpe alors de domaine de définition de cette variable en intervalles de largeur $d\theta$.

A chaque valeur comprise entre θ et $(\theta + d\theta)$ correspond une cause possible C_i de l'événement E . Les états de la nature, ensemble des causes C_i , sont représentés suivant leur importance par une densité de probabilité continue de la variable cause Θ : $f_0(\theta)$ qui n'est autre que l'histogramme de variable Θ .

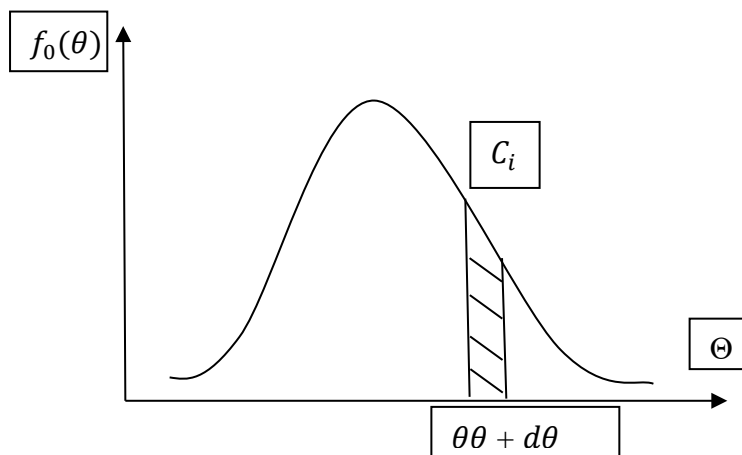


Figure 22 Variable continue

Ainsi la probabilité $P(C_i)$ pour que la variable soit comprise entre θ et $(\theta + d\theta)$ est donnée par : $P(E|C_i) = f_0(\theta) \cdot d\theta$

On obtient ainsi, à partir de la distribution a priori $f_0(\theta)$ et de l'événement objectif E qui est par exemple le résultat d'essais :

$$P(E|C_1) = f_1(E|\theta)$$

Avec f_1 la fonction de vraisemblance qui représente par exemple la probabilité d'obtenir K défaillances sur un temps cumulé T sachant que la variable est $\Theta = \lambda$.

On en déduit la probabilité a posteriori que le paramètre cause soit comprise entre θ et $(\theta + d\theta)$ sachant que l'événement E a eu lieu, et on retrouve le théorème de Bayes tel qu'il a été défini précédemment.

$$f_2(\theta / E)d\theta = \frac{f_1(E / \theta)f_0(\theta)d\theta}{\int_{D(\theta)} f_1(E / \theta)f_0(\theta) d\theta}$$

$f_2(\theta | E)$ est la densité de probabilité a posteriori de θ .

$D(\theta)$ est le domaine de définition de la variable aléatoire cause qui intervient dans l'a priori. Il est en général fixé par les experts. On obtient alors une nouvelle répartition du paramètre θ tenant compte des nouvelles informations acquises et de la distribution a priori.

5. L'approche bayésienne adoptée en pratique

L'approche bayésienne adoptée en pratique n'utilise pas des lois de probabilité conditionnelles mais des espérances conditionnelles. Elle traite des modèles d'évolution seulement en moyennes quadratiques.

L'approche dynamique de l'analyse bayésienne usuelle est basée sur l'arrivée au cours du temps des données issues du retour d'expériences. Cette approche peut se concevoir à travers une évolution temporelle avec une loi a priori au début du processus et la loi a posteriori à la fin du processus.

Attention, le temps intervient dans l'inférence bayésienne sous la forme des données d'essais ou d'expérience qui ont été réalisés à un instant donnée.

Ici, les données d'expériences arrivent au cours du temps et la connaissance instantanée de la fiabilité est réactualisée. Le retour d'expérience apparaît, soit par des observations sur le fonctionnement des composants, soit par des lois de comportement de composant antérieurs, pour alimenter un modèle dynamique d'évolution du taux de défaillance.

L'application de la théorie de Bayes en fiabilité utilise la notion de loi de probabilité a priori et de loi de probabilité a posteriori, dont les paramètres sont estimés par la méthode du maximum de vraisemblance. Or, souvent, les données disponibles ne concernent que les deux premiers moments, parfois même un seul chiffre de fiabilité. Le problème est d'extraire le maximum d'information de ses données sans créer artificiellement de l'information.

Cette approche modélise l'information sous la forme de lois de probabilité paramétrique et c'est la valeur du ou des paramètres qui est à déterminer. On utilise pour cela une méthode de maximum de vraisemblance a posteriori quand les premières données sont disponibles.

Puis, quand celles-ci sont en nombre suffisant, on peut construire la loi de probabilité a posteriori.

On dispose donc au cours du temps, dans leur ordre de détermination de :

- la loi de probabilité a priori.
- la loi conditionnelle des observations recueillies ou fonction de vraisemblance.
- la loi de probabilité a posteriori.
- la loi de probabilité du maximum de vraisemblance conditionnel a posteriori.

Cette approche ne répond pas seulement à la demande de prédiction de paramètre de fiabilité dans le futur à partir de la connaissance présente et passée, mais aussi à la nécessité de définir les paramètres d'un système dès sa conception.

Le principe de l'inférence bayésienne est le suivant. Soit la distribution d'une variable aléatoire, de vecteur paramètre θ , d'après le théorème de Bayes, on peut écrire :

$$P(\theta | Donnée) = P(Donnée | \theta).P(\theta)$$

Les sources d'information a priori peuvent être objectives ou subjectives. Elles sont illustrées sur la figure suivante.

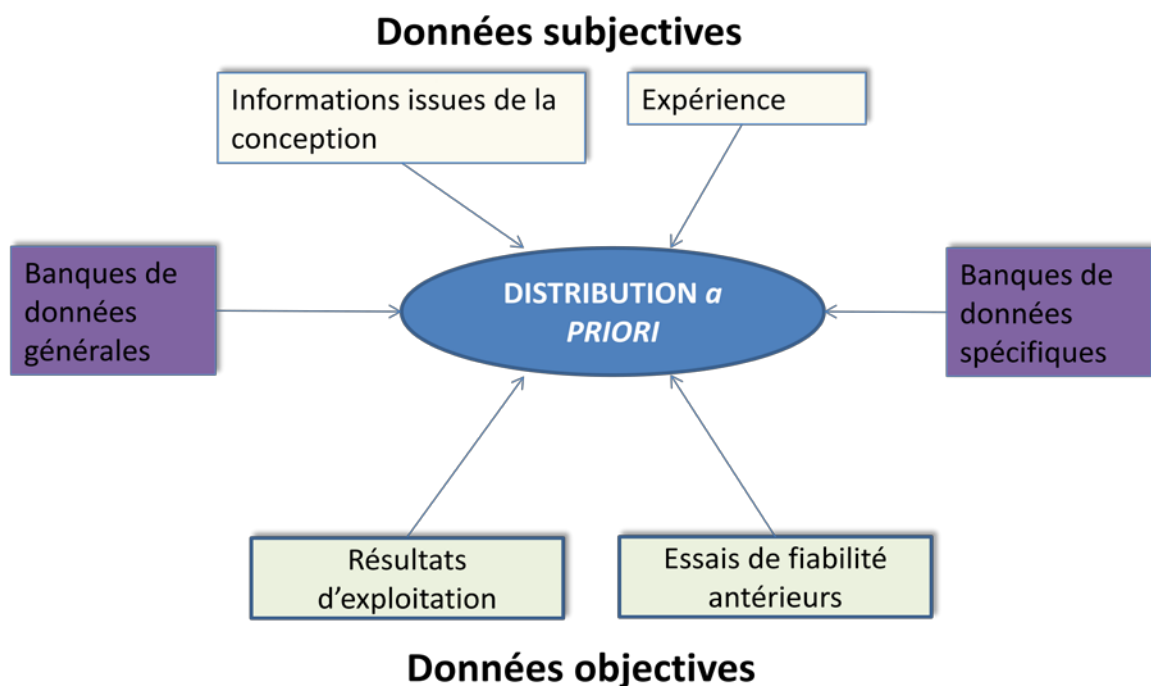


Figure 23 Les données à priori

6. Le mécanisme d'inférence

Le processus d'inférence comporte les étapes suivantes :

- choix d'un modèle relatif aux données : distribution conditionnelle. La distribution conditionnelle appartient à une famille de lois de probabilité.
- détermination d'une distribution à priori du paramètre de la loi conditionnelle, pour intégrer des informations subjectives. Elle est établie à partir des connaissances que l'on a sur le matériel étudié : essais antérieur sur le matériel similaire, avis d'experts, retour d'expérience, enquête, etc.
- obtention de la distribution a postériori par application du théorème de Bayes qui permet de combiner l'information subjective et l'information objective contenue dans l'échantillon.
- estimation du paramètre de fiabilité étudiée (taux de défaillance par exemple) déduite de la distribution a posteriori.

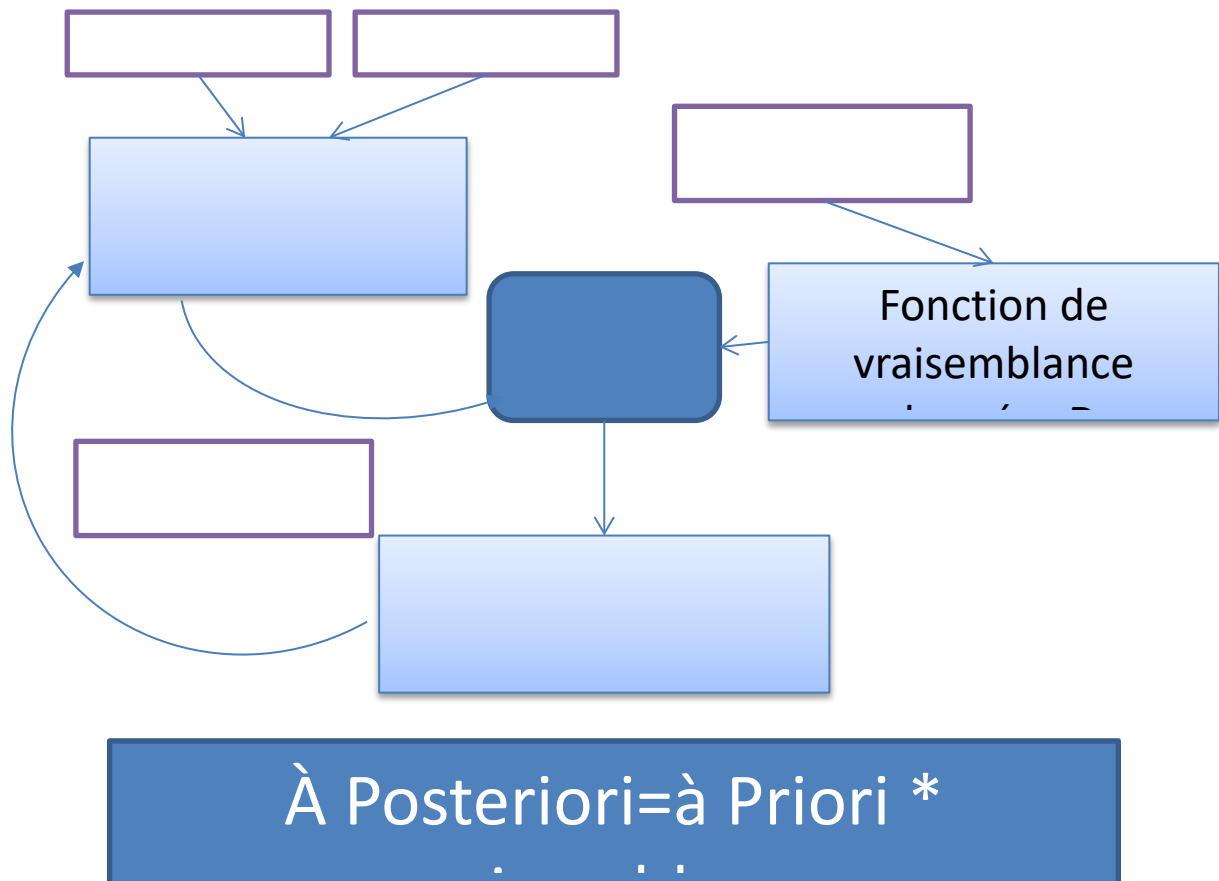


Figure 24 Mécanisme d'inférence

Dans le prochain chapitre, nous détaillerons le côté pratique de cette approche et nous présenterons aussi un outil très connu qui est BayesiaLab.

CHAPITRE 5 : Réseaux bayésiens et l'outil « BayesiaLab »

1. Introduction :

Les premières formalisations du Réseau Bayésien ont été initiées par les travaux de PEARL [14] pour mettre en évidence trois aspects :

- la nature subjective des informations.
- l'utilisation de la règle de Bayes comme principe de base pour la mise à jour des informations.
- la distinction entre les modes de raisonnement causal et fondée.

Les approches fondées sur le conditionnement et la formule d'inversion de Bayes sont appelées approches bayésiennes. Le nom de Réseaux Bayésiens est donné par cohérence avec cette appellation. D'autres noms existent également pour ces modèles : les réseaux de croyance, les modèles graphiques probabilistes orientés, les réseaux probabilistes, ou encore les réseaux d'indépendances probabilistes.

Le réseau bayésien est un système représentant la connaissance et permettant de calculer des probabilités conditionnelles apportant des solutions à différentes sortes de problématiques.

Ce réseau a été initié tout d'abord dans le domaine de l'intelligence artificielle, puis élargi dans les autres domaines scientifiques, en particulier dans le domaine de sûreté de fonctionnement pour la gestion des événements incertains, la détection et le diagnostic des défaillances des systèmes de procédés..... etc.

Les recherches et développements apportés aux Réseaux Bayésiens au cours de ces dernières années, portent principalement sur trois points essentiels : les algorithmes d'inférence, l'apprentissage de la structure du réseau et l'apprentissage des paramètres du réseau. Ces algorithmes sont opérationnels dans le cas des Réseaux Bayésiens classiques (statiques).

2. Structure du Réseau Bayésien (RB)

Un réseau bayésien est un outil complet permettant la visualisation de variables et de leurs dépendances ou indépendances. Il permet également de décrire quantitativement le fonctionnement d'un système grâce aux différents calculs de probabilités concernant les variables du système.

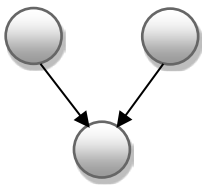
Ce réseau est un graphe dans lequel les nœuds représentent des variables aléatoires discrètes ou continues, et les arcs reliant ces dernières sont rattachées à des probabilités conditionnelles.

Le fait d'indiquer un arc entre deux variables implique une dépendance directe entre ces deux variables : l'une est le parent et l'autre l'enfant. Il faut fournir le comportement de la variable enfant à partir du comportement de son ou ses parents.

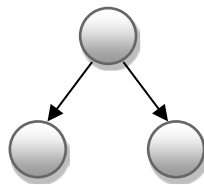
Pour cela, chaque nœud du réseau possède une table de probabilités conditionnelles. Cette dernière associée à un nœud permet de quantifier l'effet du ou des nœuds parents sur ce nœud, elle décrit les probabilités associées aux nœuds enfants suivant les différentes valeurs des nœuds parents. Pour les nœuds racines (sans parents), la table de probabilité n'est plus conditionnelle et mais fixe alors des probabilités a priori concernant les valeurs de la ces variables doit être définie.

Le graphe est acyclique : il ne contient pas de boucles. Les arcs représentent des relations entre variables qui sont soit déterministes, soit probabilistes. Ainsi, l'observation d'une ou plusieurs causes n'entraîne pas systématiquement l'effet ou les effets qui en dépendent, mais modifie seulement la probabilité de les observer.

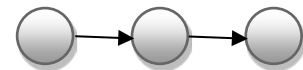
- Plusieurs types de chaînes sont possibles entre 3 variables :



Connexion convergente



connexion divergente



connexion normale

De manière formelle, un Réseau Bayésien est défini par :

- un graphe acyclique orienté $G, G = (V, E)$, où V est l'ensemble des nœuds de G , et E est l'ensemble des arcs de G .

- un espace probabilisé (Ω, Z, P) , avec Ω un ensemble fini non vide, Z un ensemble de sous-espaces de Ω , et P une mesure de probabilité sur Z avec $P(\Omega) = 1$.

- un ensemble de variables aléatoires associées aux nœuds du graphe G et défini sur

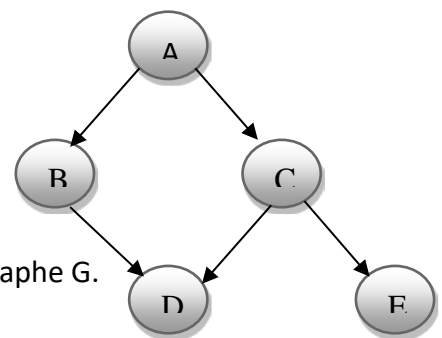
(Ω, Z, P) , tel que :

$$P(V_1, V_2, \dots, V_n) = \prod_{i=1}^n P(V_i / CV_i)$$

Où $C(V_i)$ est l'ensemble des parents (ou causes) de V_i dans le graphe G .

Exemple illustratif :

$$P(A,B,C,D,E) = P(D/BC) P(E/C)P(B/A)P(C/A)P(A)$$



L'intérêt particulier des réseaux bayésiens est de tenir compte, simultanément, de connaissances a priori d'experts (dans le graphe) et de l'expérience contenue dans les données.

Un autre intérêt de l'utilisation des Réseaux Bayésiens dans l'analyse des événements incertains, réside dans la possibilité d'effectuer une description qualitative entre les différentes variables représentant ces événements (le graphe causal) et la description quantitative des dépendances entre les événements (probabilités conditionnelles).

La construction d'un Réseau Bayésien passe par trois étapes primordiales :

- La structure : comment déterminer la structure du réseau ou le graphe causal.
- L'inférence : comment calculer les probabilités a posteriori de chacune des variables du réseau. Généralement, l'inférence est lancée dès qu'une information nouvelle concernant une ou plusieurs variables est disponible. Cet apport d'information est appelé évidence et elle est propagée dans le réseau par le moteur d'inférence.
- L'apprentissage : comment estimer les probabilités et la structure.

2.1 Tables de probabilités

Les tables de probabilités, liées aux nœuds, sont définies par des statistiques relatives au problème à résoudre (peuvent aussi être déterminées par des experts). Chacune des variables dispose d'une table de probabilités conditionnelles relatives aux variables causales dont elle dépend.

2.2 Notion de causalité et relations entre les nœuds

Dans les Réseaux Bayésiens, deux nœuds qui représentent deux faits différents peuvent être en relation causale sans que l'un implique l'autre. Graphiquement, chaque parent a un effet sur ses fils. La notion de causalité joue un rôle très important pour construire les Réseaux Bayésiens.

Un Réseau Bayésien permet de modéliser plusieurs types de nœuds. Nous utilisons dans la plupart des cas deux types de nœuds :

- un nœud discret représentant une variable discrète.
- un nœud continu représentant une variable continue.

2.3 Cas de deux variables discrètes

En dressant un arc allant de A vers B, qui sont variables discrètes multinomiales de dimension respective a et b avec a_1, a_2, \dots, a_a les différentes modalités de A, et b_1, b_2, \dots, b_b les différentes modalités de B, la table de probabilités conditionnelles se présente comme suit :

Tableau 8 table de probabilités conditionnelles à deux variables discrètes

A	B			
	b_1	b_2	...	b_b
a_1	$P(b_1/a_1)$	$P(b_2/a_1)$...	$P(b_b/a_1)$
a_2	$P(b_1/a_2)$	$P(b_2/a_2)$...	$P(b_b/a_2)$
⋮	⋮	⋮	⋮	⋮
a_a	$P(b_1/a_a)$	$P(b_2/a_a)$...	$P(b_b/a_a)$

Cette table de probabilités conditionnelles permet d’avoir toutes les informations nécessaires à l’inférence dans un réseau. La taille de cette table est de $a \times b$ et peut devenir très grande pour des variables avec beaucoup de modalités, donc son remplissage peut devenir difficile. En effet, la taille d’une table de probabilités d’un nœud discret X de taille x , ayant p parents discrets Y_1, Y_2, \dots, Y_p de tailles respectives y_1, y_2, \dots, y_p est de $X * \prod_{i=1}^p y_i$

2.4 Cas de deux variables discrète et continue

Dans le cas de deux variables, A discrète multinomiale de dimension a et B continue de paramètre μ_B et Σ_B . En dressant un arc partant de A vers B , la table de probabilités conditionnelles de B se présente comme suit :

Tableau 9 table de probabilité à variables discrètes et continues

A	B
a_1	$P(B/a_1) = N(\mu_{a_1}, \Sigma_{a_1})$
a_2	$P(B/a_2) = N(\mu_{a_2}, \Sigma_{a_2})$
...	...
a_a	$P(B/a_a) = N(\mu_{a_a}, \Sigma_{a_a})$

La table de probabilité de B se compose de lois conditionnées aux modalités de A . En effet, la table de probabilité d’un nœud continu X , ayant p parents discrets Y_1, Y_2, \dots, Y_p de tailles respectives y_1, y_2, \dots, y_p est de $X * \prod_{i=1}^p y_i$ lois continues.

2.5 Cas de deux variables continues

Dans le cas de deux variables continues A et B de paramètres respectifs μ_A, Σ_A et μ_B, Σ_B . En dressant un arc partant de A vers B , on effectue alors une régression et l’on peut écrire la loi régissant B pour une valeur à a de A comme étant une loi gaussienne de paramètres

$(\mu_B + \beta \times A ; \Sigma_B)$, ou β représente le coefficient de régression.

3. Calculs d’inférences dans un Réseau Bayésien

Les calculs dans un réseau bayésien se nomment inférence. L’inférence permet de calculer les probabilités a posteriori de chacune des variables du réseau.

Généralement, l'inférence est lancée dès qu'une information nouvelle concernant une ou plusieurs variables est disponible. Cet apport d'information est appelé évidence. Une fois l'information indiquée, celle-ci est propagée dans le réseau par le moteur d'inférence.

Les algorithmes les plus utilisés pour le calcul exact des inférences dans le cas de Réseaux Bayésiens avec variables discrètes sont: « Pearl Message Passing Algorithm » : cet algorithme a été développé par Pearl en 86 [14] en exploitant les indépendances locales des événements du Réseau Bayésien. Étant donné un ensemble de valeurs a d'un ensemble A de variables, l'algorithme détermine $P(x|a)$ pour toutes les valeurs x de chaque variable X du réseau. Il accomplit cette mission en initiant des messages de chaque variable instanciée pour ses voisins. Ces voisins à leur tour, transmettent des messages à leurs voisins. Cette méthode ne fonctionne de manière exacte que lorsque le Réseau Bayésien possède une structure sous forme d'arbre [15].

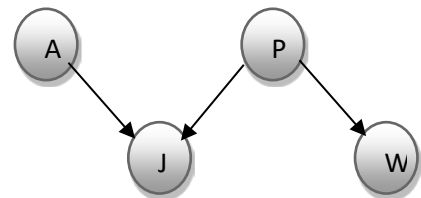
4. Exemple d'un réseau bayésien

Ce matin-là, alors que le temps est clair et sec, M. Riadh sort de sa maison. Il s'aperçoit que la pelouse de son jardin est humide. Il se demande alors s'il a plu pendant la nuit, ou s'il a simplement oublié de débrancher son arroseur automatique.

Il jette alors un coup d'œil à la pelouse de son voisin, M. Amine, et s'aperçoit qu'elle est également humide. Il en déduit alors qu'il a probablement plu, et il décide de partir au travail sans vérifier son arroseur automatique.

Prenons :

- A : J'ai oublié de débrancher mon arroseur automatique.
- P : Il a plu pendant la nuit.
- J : L'herbe de mon jardin est humide.
- W : L'herbe du jardin de M. Amine est humide.



Arbre des probabilités à Priori

Tableau 10 Probabilité a posteriori

Événement	Probabilité	Commentaire
A=V	0,4	M. Riadh oublie assez souvent de débrancher son arroseur automatique.
A=F	0,6	
P=V	0,6	La région est relativement pluvieuse
P=F	0,4	

	A=V		A=F	
	P=V	P=F	P=V	P=F
J=V	1	1	1	0
J=F	0	0	0	1

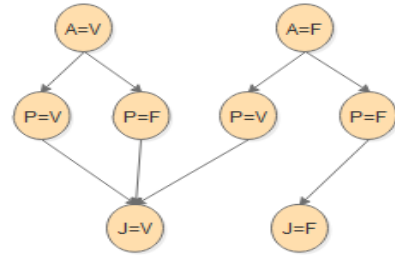


Figure 25 Arbre de probabilité

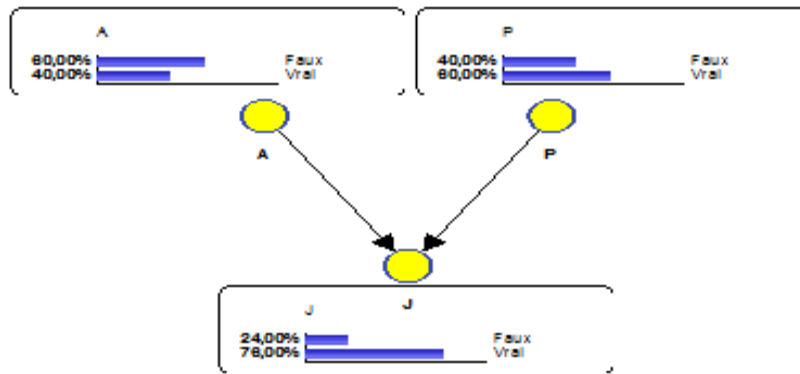


Figure 26 Réseau bayésien par BayesiaLab 1

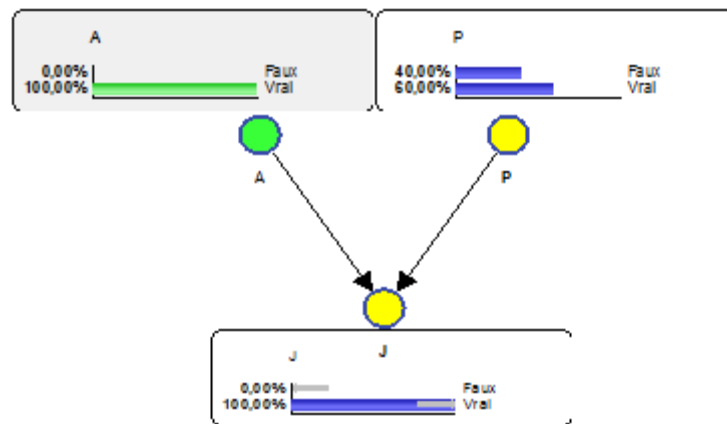


Figure 27 Réseau bayésien par BayesiaLab 2

(M. Riadh a oublié de débrancher son arroseur automatique).

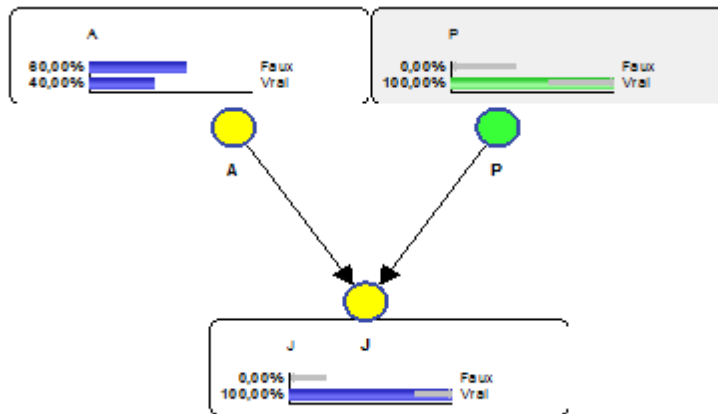


Figure 28 Réseau bayésien par BayesiaLab (Il a plu pendant la nuit).

5. Structure du réseau bayésien à partir de l'arbre des défaillances " conditions pour une mutation de l'AdD au RB"

Dans le cas de quelques études particulières, des algorithmes ont été développés pour le passage de la structure des arbres des défaillances aux réseaux bayésiens lorsque les hypothèses de base de la construction de l'arbre de faute sont satisfaites :

- Les événements sont binaires (fonctionne / ne fonctionne pas).
- Les événements sont statistiquement indépendants.
- Les relations entre les événements et les causes sont représentées par des portes logiques ET et OU.
- La cible de l'arbre de fautes est l'événement indésirable qu'il faut analyser [15].

L'algorithme de transformation de la structure (AdD) à celle du (RB) comprend donc des tâches graphiques et d'autres numériques. Dans la transformation de la structure graphique, les événements primaires, les événements intermédiaires et les événements sommets de l'arbre des défaillances sont représentés respectivement comme des nœuds racines, des nœuds intermédiaires, et des nœuds terminaux dans la structure des réseaux bayésiens équivalents.

Les nœuds du réseau bayésien équivalent sont connectés de la même manière que leurs correspondants composants dans l'arbre des défaillances. La figure suivante illustre la procédure simplifiée de transformation du modèle arbre des défaillances en réseau bayésien.

Dans la transformation numérique, les probabilités d'occurrence des événements primaires sont affectées aux nœuds racines correspondants comme des probabilités a priori. Pour chaque nœud intermédiaire et terminal du graphe, une table de probabilité conditionnelle (TPC) est développée selon le type des portes logiques.

La figure suivante montre la conversion d'une porte logique OU et ET en nœuds équivalents dans le RB. Les nœuds parents A et B sont affectés de probabilités a priori, et au nœud enfant C sa table de probabilité conditionnelle (TPC).

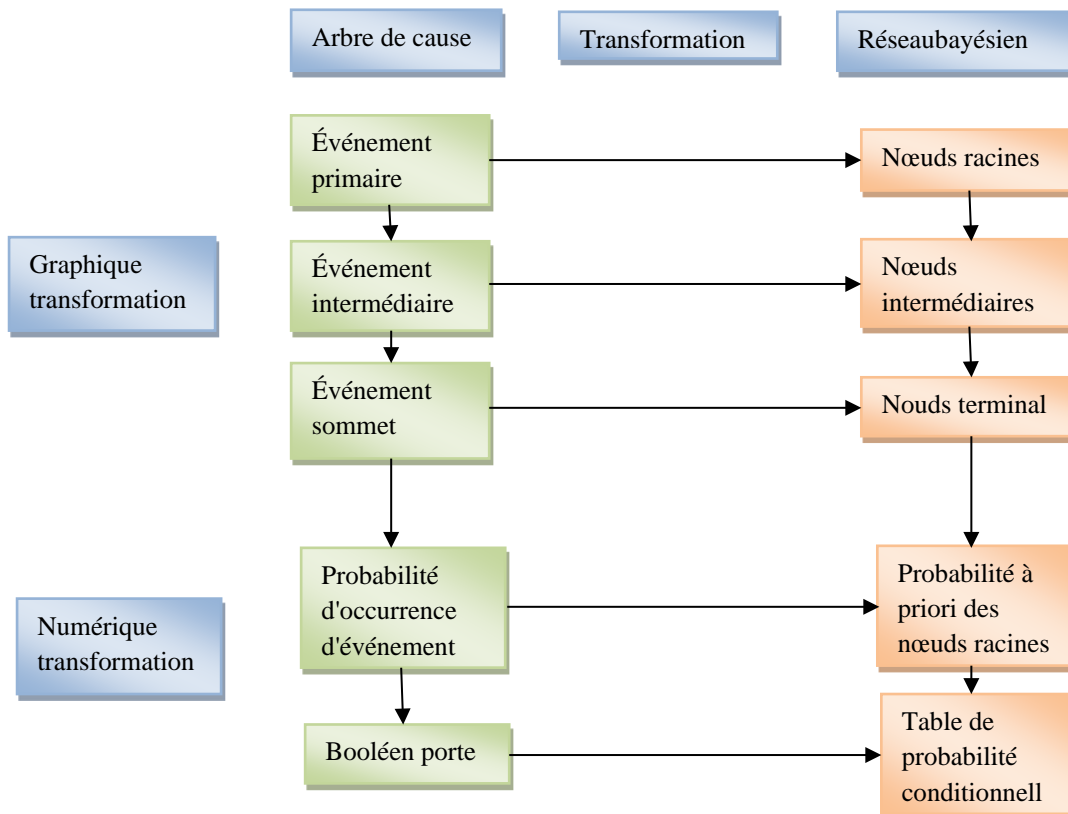


Figure 29 Transformation depuis l'arbre des défaillances au réseau bayésien

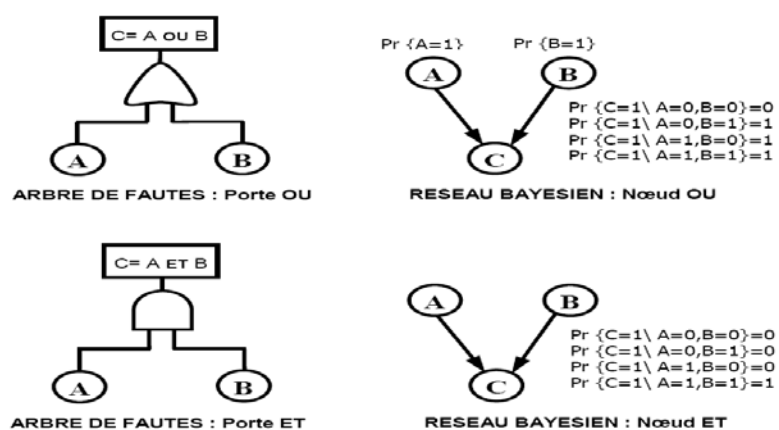


Figure 30 Figure 1. Présentation des portes logiques " OU et ET " dans l'Add et le RB

6. Introduction sur BayesiaLab

Le développement théorique des réseaux bayésiens date des années 1970, mais à cette époque, le manque d'algorithmes efficaces et d'outils pratiques a empêché le développement d'application.

Aujourd'hui, grâce à des outils puissants et conviviaux, les réseaux bayésiens sont une technique en pleine expansion dans beaucoup de domaines où l'on a besoin d'aide à la décision dans un contexte de connaissance incertaine du (monde réel), par exemple dans les domaines : médical, militaire, financier, robotique, météorologique, etc. BayesiaLab est l'un de ces outils.

BayesiaLab est un produit de Bayesia (<http://www.bayesia.com>), entreprise française dédiée à l'utilisation des méthodes d'aide à la décision et d'apprentissage issues de l'intelligence artificielle ainsi qu'à leurs applications opérationnelles (industrie, services, finance, etc.).

BayesiaLab se présente comme un laboratoire complet de manipulation et d'étude de réseau bayésien. Il est développé en Java, et est actuellement disponible en version française ou anglaise.

BayesiaLab permet de traiter l'ensemble de la chaîne d'étude de la modélisation d'un système par réseau bayésien : modélisation, apprentissage automatique, analyse et utilisation.

BayesiaLab est le résultat de presque vingt ans de développement de logiciel par Dr. Lionel Jouffe et Dr. Paul Munteanu. En 2001, leurs efforts de recherches menés à la création de Bayesia, siégée dans Laval en France du nord-ouest.

Aujourd'hui, la compagnie est devenue le principal fournisseur des technologies liées au réseau bayésien pour des centaines de sociétés importantes autour du monde [16].

7. Modélisation de la connaissance par BayesiaLab

BayesiaLab permet de réaliser une représentation visuelle des causes et des effets. Des nœuds (représentant des variables) peuvent être ajoutés et placés avec un clic de souris, arcs (représentant des rapports) peuvent « être dessinés » entre les nœuds. La direction causale est simplement codée dans la direction de l'arc.

La nature quantitative des dépendances, et beaucoup d'autres attributs peuvent être contrôlés dans le rédacteur de nœud, qui est disponible en cliquant par la droite sur n'importe quel nœud.

BayesiaLab facilite ainsi de coder intuitivement sa propre classification d'un domaine avec un minimum d'effort. Simultanément il impose l'uniformité interne, de sorte qu'aucune condition impossible ne soit accidentellement codée.

En plus de permettre à des utilisateurs de coder directement leur connaissance explicite en dessinant un réseau dans le panneau de graphique, l'environnement d'obtention de connaissance approfondie de Bayesia est disponible comme prolongation à BayesiaLab. Il laisse obtenir systématiquement la connaissance explicite et tacite d'un groupe d'experts pendant les sessions de séance de réflexion.

8. Exemple d'une simulation avec BayesiaLab

Supposons la situation suivante :

Un patient examiné pour une maladie rare, qui affecte aléatoirement un dans 10.000 personnes. Après les tests, le docteur donne les nouvelles d'un résultat positif, c.-à-d. Que le patient est atteint de cette maladie. Le docteur déclare également que cet essai est avéré être « 99% correct, » ce qui signifie qu'un positif est faux à 1% et un résultat négatif est faux à 1%.

Le patient sera vraisemblablement choqué et se prépare au pire. Cependant, est-ce que c'est une évaluation correcte de la situation ? Etant donné le résultat positif, est-il vraiment concerné ?

Un réseau bayésien très simple, peut être créé, pour coder l'information de fond que nous avons, puis ajouter l'évidence, et effectuer finalement l'inférence pour déterminer la probabilité vraie d'être malade.

Notre domaine de problème peut être modélisé avec deux variables binaires, une pour la maladie, c.-à-d., qu'il soit malade ou pas (vrai /faux), et les autres pour le résultat du test, c.-à-d. positif ou négatif.

Les variables sont représentées comme nœuds dans BayesiaLab. Les nœuds peuvent être ajoutés au panneau de graphique en cliquant d'abord sur l'icône de mode de création nœud et en indiquant ensuite la position désirée sur l'écran avec l'indicateur.

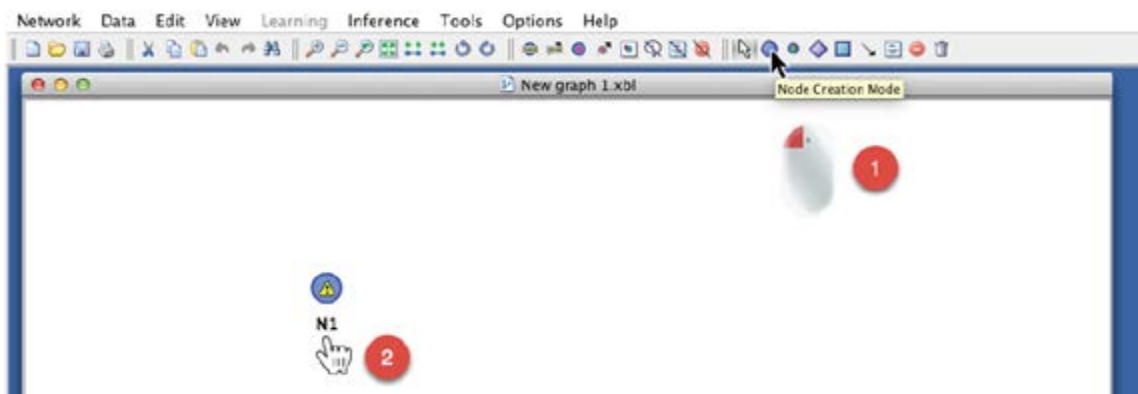


Figure 31 Création des nœuds

Sur le premier nœud, propriété, Choisir alors le premier onglet pour voir les états du nœud. Par défaut, deux états, vrai et faux, sont déjà définis, vous pouvez garder ainsi pour la maladie.

Dans l'onglet suivant, (distribution de probabilité), on clique sur " probabiliste ".

C'est là où le premier bout de la connaissance à priori peut être introduite, qui n'est autre que la probabilité de la présence de la maladie qui est 1 pour 10.000 ou 0.01%.

Écrire la valeur de pourcentage en colonne vraie, puis cliquer dans la cellule fausse, et puis cliquer compléter.

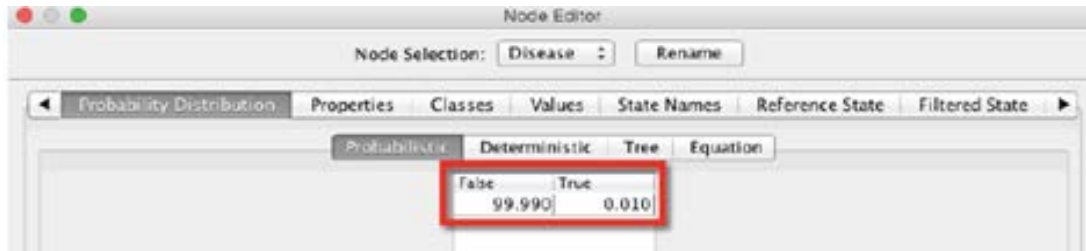


Figure 32 Propriété du nœud maladie

Cependant, pour le Test, nous ne pouvons pas simplement définir les probabilités car ce n'est pas une distribution marginale comme nous avons fait pour la maladie. En revanche, nous devons définir ces probabilités en fonction des états de la maladie.

Une « fonction » est représentée par un arc dirigé dans un réseau bayésien. Ainsi, vous devez créer un arc de la maladie vers le Test. Cliquer sur l'icône de mode de création d'arc, cliquer sur la maladie, et traîner le curseur jusqu'au Test, et puis libérer.

Dans les propriétés du nœud test de nouveau, les deux valeurs par défaut sont vraies et faux. Dans le cadre de cet exemple, il serait plus approprié de se rapporter à un résultat de Test positif ou négatif. Double clic dans la liste pour changer vrai et faux en positif et négatif respectivement.

Après, dans l'onglet de probabilité (probability distribution), une table blanche avec les états de Test en colonne les états de la maladie en lignes.

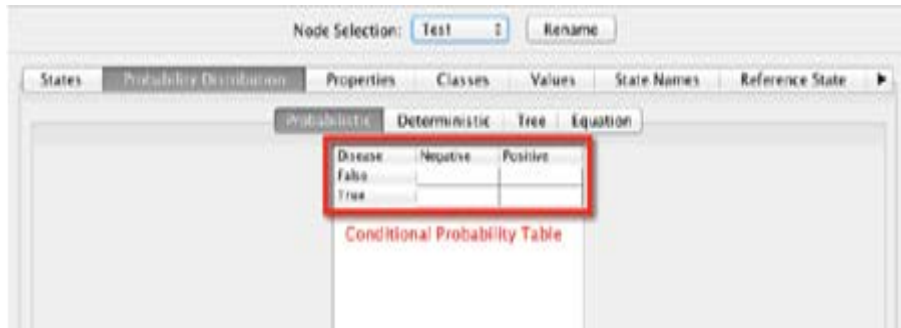
Dans le cadre de réseau bayésien, ceci est connu comme Tableau conditionnel de probabilité. Il définit les probabilités des états du nœud étant donné son nœud parent (ou nœuds parent). Etant donné la direction d'arc de maladie vers Test, l'origine est le nœud parent et l'arrivée est le nœud enfant.

Puisque la maladie n'a aucun nœud parent, nous avons seulement écrit des probabilités marginales. Cependant, le nœud Test a la maladie comme nœud parent, et nous maintenant devons écrire des probabilités conditionnelles.

Du rapport du médecin nous savons que le Test est 99% correct. Ceci signifie que si la maladie était vraiment présente chez le patient, le test donnerait un résultat correct et positif 99% du temps. Il signifie également que le test produirait inexactement un résultat négatif dans 1% des cas. Dans cet exemple d'introduction, nous supposons plus loin que « 99% correct

» s'applique également à l'absence de la maladie, c.-à-d. le test indiquerait un résultat positif dans 1% des cas de patients sains.

Écrire ces valeurs dans le Tableau de probabilité conditionnelle (Conditional Probability Table) et cliquez sur accepter (accept).



Test		
Disease	Negative	Positive
False	99.000	1.000
True	1.000	99.000

Figure 33 Table des probabilités du nœud test

Que pouvons-nous calculer avec ce réseau ? Nous pouvons effectuer l'inférence dans ce réseau. Ceci signifie que nous pouvons placer l'évidence sur un nœud et puis donner lecture de la distribution de probabilité de l'autre nœud. Vous pouvez objecter et dire que nous savons déjà ça, c.-à-d. étant donné que la maladie est présente, il y a une probabilité de 99% que le test sera positif. En fait, c'est déjà une donnée du réseau. Cependant, le calcul n'est plus insignifiant si nous plaçons l'évidence sur le test à la place, car nous essayons de raisonner dans la direction opposée, du test vers la maladie, qui était la motivation pour cet exercice.

Pour exécuter l'inférence dans BayesiaLab, il faut commuter du mode de modélisation (Modeling Mode), au mode de validation (Validation Mode). Nous pouvons commuter dans les deux sens entre ces deux modes par l'intermédiaire des icônes dans le coin bas-gauche du panneau de graphique :



Figure 34 Les deux modes

Maintenant que nous sommes en mode de validation, une nouvelle section apparaît à la droite du panneau de graphique. C'est le panneau de moniteur.

En double cliquant sur les deux nœuds, maladie et Test, leurs modalités et leurs valeurs apparaissent sur le Monitor.

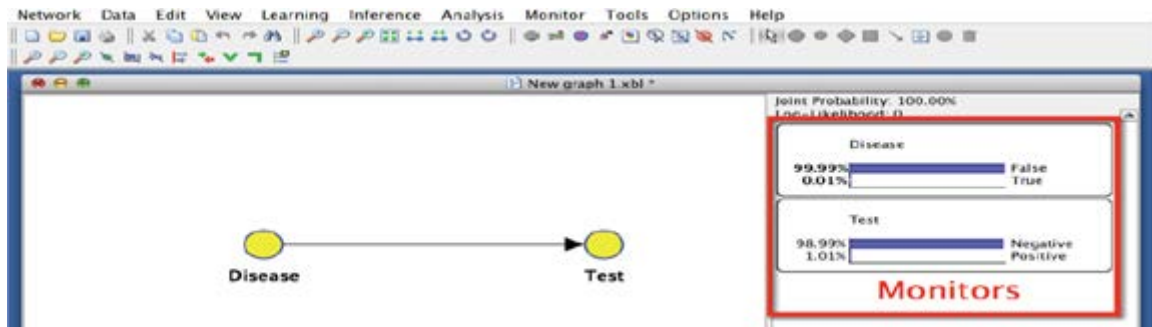


Figure 35 Les modalités

Les moniteurs montrent toujours la distribution de probabilité courante des nœuds. Ils permettent également de manipuler des probabilités. Dans la plupart des cas, ceci signifie qu'une actualisation du réseau a été effectuée en modifiant une probabilité afin de représenter une évidence constatée par l'observation.

Nous voulons maintenant déduire la probabilité de la maladie, sachant que le résultat du test est positif. Ce type d'inférence est également connu dans le diagnostique.

Ainsi, sur le moniteur Test à Positive est mise à =100%.

Ce qui est vraiment intéressant, c'est de voir le changement du moniteur de la maladie. En changeant les valeurs de Test, BayesiaLab a automatiquement calculé la nouvelle probabilité pour la maladie en utilisant la règle de Bayes. Cependant, en regardant le moniteur de maladie, on se demande si quelque chose avait changé. Au premier regard, la longueur de la barre pour faux semble pratiquement inchangée.

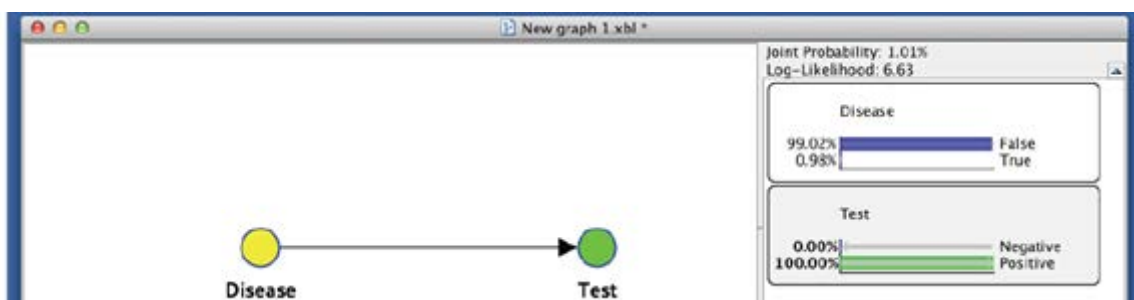


Figure 36 Changement des modalités

En fait oui. La probabilité marginale d'avoir la maladie était à l'origine 0.01% (une dans 10.000, comme elle est pour tout le monde). Maintenant, étant donné le résultat du test positif, elle est 0.98%, c.-à-d. la probabilité d'avoir la maladie a augmenté 98 fois.

Ainsi, l'évidence a rendu la maladie pour ce patient beaucoup plus probable qu'avant, cependant, c'est nullement une preuve (car ça reste une probabilité). En revanche, il y a toujours une chance de 0,98% que le patient soit en bonne santé.

Des exemples détaillés seront donnée dans le support de travaux dirigés.

CHAPITRE 6 : Modélisation des données et principe de conjugaison

1. Introduction

La démarche bayésienne a été formalisée et relancée au 20ème siècle dans les années 1930 [17]. Les difficultés algorithmiques sont maintenant résolues grâce au développement du calcul intégral puis de logiciels dédiés et à l'emploi des ordinateurs [17]. Mais les difficultés majeures liées à la démarche bayésienne restent :

Le choix d'une distribution de probabilité a priori appropriée à l'état des connaissances initiales, est une démarche complexe, qui peut laisser place à une appréciation subjective dont il est difficile de spécifier une loi a priori précise pour les modéliser. Malgré latitude dans le choix de la loi a priori, on souhaite lui associer les propriétés suivantes :

- Le calcul de la densité jointe a posteriori à partir du produit des distributions a priori et celle de l'échantillon d'observation doit être aussi simple que possible ;
- La distribution a posteriori doit être de préférence du même type que la distribution a priori afin de permettre un calcul d'actualisation itératif ;

Dans le domaine de la fiabilité des systèmes et pour des raisons de conjugaison des lois (voir section suivante), les principales distributions utilisées sont :

- La loi exponentielle qui est directement associée au processus poissonnien et qui est représentative de la durée de vie de matériels qui ont un taux de défaillance aléatoire constant,
 - La loi de Weibull qui généralise la loi exponentielle mais dont le taux de défaillance varie dans le temps,
 - La loi Log-Normale représentative en particulier des temps de réparation des matériels et enfin la loi Gamma.
- La distribution a priori doit pouvoir représenter un grand nombre de situations ou de phénomènes physiques ;
 - Elle doit être paramétrable, et les paramètres doivent pouvoir être interprétés physiquement.

- L'informativité des connaissances a priori a un impact sur la densité a posteriori (résultat). La probabilité a posteriori est une moyenne pondérée entre l'information fournie a priori et les observations du retour d'expérience (la fonction de vraisemblance) : intuitivement elle sera attirée par la distribution qui apportera le plus d'informations, donc celle qui sera la plus précise par rapport à sa valeur moyenne. Pour la fiabilité des systèmes mécaniques, les données disponibles avant (a priori) la collecte d'un retour d'expérience (la fonction de vraisemblance) sont principalement de deux formes:

- informations objectives: Ce sont des données opérationnelles, issues d'essais ou de retour d'expérience (REX) auprès des fabricants ou des recueils de données connues.
- informations subjectives: Ce sont des données représentant l'avis d'un ou de plusieurs experts.

Ainsi, en fonction du type de distribution choisie pour modéliser l'information a priori qui peut être plus ou moins informative vis-à-vis des observations, donc plus ou moins d'impact sur la densité de probabilité a posteriori.

2. Principe de conjugaison

Le principe de conjugaison utilisé dans la démarche Bayésienne consiste à choisir une loi a priori de même nature que la fonction de vraisemblance afin d'aboutir à une distribution a posteriori connue. En effet, si la densité a priori est conjuguée à la fonction de vraisemblance, la détermination de la distribution jointe a posteriori, est simplifiée, puisque ses paramètres sont tout simplement égaux à la somme des paramètres de la densité a priori et de la fonction de vraisemblance.

Le produit de deux distributions conjuguées est une distribution elle-même conjuguée aux deux autres. Dans ces conditions, le calcul ne présente aucune difficulté et peut se réaliser manuellement.

Il est démontré que les familles de lois dites conjuguées ne se contentent pas uniquement de simplifier les calculs mais surtout que ce type de lois que l'on peut le mieux maîtriser pour pondérer à volonté le pouvoir relatif entre les données et la connaissance a priori dans la détermination de la loi a posteriori. On peut ainsi minimiser l'impact de la loi a priori si le REX est fiable, ou au contraire le favoriser dans le cas où le système a subi des modifications [17]. L'impact des modifications est alors estimé par l'expertise a priori.

Le tableau suivant présente les principales distributions conjuguées de lois fréquemment utilisées en fiabilité.

Tableau 11 Principales lois conjuguées

Vraisemblance Lois des observations (X/θ)	Distribution a priori (θ)	Distribution a posteriori (θX)
Poisson (θ),	Gamma (α),	Gamma ($\alpha+x,+1$)
Exponentielle (k,λ)	Gamma (α),	Gamma ($\alpha+k,+t$)
Gamma (k),	Gamma (α),	Gamma ($\alpha+k,+t$)
Binomial ($a,,$)	Bêta (α),	Bêta ($\alpha+a,+n-a$)
Bêta (α),	Exponentielle (λ)	Exponentielle ($\lambda-\ln(1-x)$)
normale	normale	normale
normale	gamma	gamma

Comme évoqué précédemment, une des difficultés de l'approche bayésienne est le calcul de la loi a posteriori. Ce calcul est facilité lorsque loi a priori et loi a posteriori ont la même forme. Dans ce cas, on parle de loi a priori conjuguée (tableau précédent). Dans ce qui suit, les techniques d'estimation sont présentées utilisant une loi exponentielle.

3. Fonction de vraisemblance Exponentielle

La loi exponentielle est la plus utilisée dans les études de fiabilité pour sa facilité d'emploi et son interprétation très simple. Elle représente la distribution de la durée de vie moyenne T d'un système, dont le taux de défaillance λ est constant avec le temps.

En mécanique, on considère que les temps de défaillance de matériels en fonctionnement suivent une loi exponentielle : dans la phase de conception (le taux de défaillance doit être prévu constant pendant la durée de mission du système), ou dans la phase d'exploitation (la maintenance a pour but de maintenir le taux de défaillance constant).

Soit la fonction de fiabilité au temps t , $R(t)$, étant :

$$R(t) = e^{-\lambda t}$$

On cherche à estimer le taux de défaillance λ sachant que t suit une loi exponentielle de paramètre λ . Sa densité de probabilité est:

$$f\left(\frac{t}{\lambda}\right) = \lambda \cdot e^{-\lambda t}$$

Lorsque les experts disposent d'un certain degré de connaissance sur le matériel considéré, il faudra choisir pour distribution a priori une loi non uniforme. Dans ce cas, la modélisation de la connaissance a priori sera par une loi informative favorisant la valeur la plus probable du paramètre considéré. Sachant que le retour d'expérience est de type exponentiel, la loi a priori la plus convenable est la loi gamma. Ces deux lois appartiennent à la même famille mathématique. Dans ces conditions, les calculs bayésiens ne présentent aucune difficulté grâce au principe de conjugaison, le produit de la loi gamma et de la loi exponentielle du REX appelée fonction de vraisemblance est une densité a posteriori qui est loi gamma.

3.1 Distribution a priori Gamma

Lorsque les connaissances disponibles sont informatives, le taux de défaillance a priori est modélisée par une distribution Gamma de paramètre α_0 et β_0 , soit :

$$g(\lambda; \alpha_0, \beta_0) = \frac{1}{\beta_0^{\alpha_0} \cdot \Gamma(\alpha_0)} \cdot \lambda^{\alpha_0-1} \cdot e^{-\frac{\lambda}{\beta_0}}$$

3.2 Fonction de vraisemblance

La fonction de vraisemblance si l'on observe k défaillances au cours d'un temps cumulé t , est déterminée de la même manière précédente, c'est-à-dire à partir de l'expression :

$$f(k/\lambda) = \lambda^k \cdot e^{-\lambda t}$$

3.3 Densité a postériori

On remplace (λ) et la vraisemblance par leur expression, l'obtention de la distribution a postériori de λ compte tenu du principe de conjugaison et de l'application du théorème de bayes est alors :

$$g(\lambda; k, \alpha_0 \beta_0) = \frac{\frac{1}{\beta_0^{\alpha_0} \Gamma(\alpha_0)} \cdot \lambda^{\alpha_0-1} \cdot e^{-\frac{\lambda}{\beta_0}} \cdot \lambda^k \cdot e^{-\lambda t}}{\int_0^{+\infty} \frac{1}{\beta_0^{\alpha_0} \Gamma(\alpha_0)} \cdot \lambda^{\alpha_0-1} \cdot e^{-\frac{\lambda}{\beta_0}} \cdot \lambda^k \cdot e^{-\lambda t} d\lambda}$$

Soit après simplification :

$$g(\lambda; k, \alpha_0 \beta_0) = \frac{\lambda^{\alpha_0+k-1} \cdot e^{-\lambda(\frac{1}{\beta_0}+t)}}{\int_0^{+\infty} \lambda^{\alpha_0+k-1} \cdot e^{-\lambda(\frac{1}{\beta_0}+t)} d\lambda}$$

Posons :

$$I = \int_0^{+\infty} \lambda^{\alpha_0+k-1} \cdot e^{-\lambda(\frac{1}{\beta_0}+t)} d\lambda \quad \text{et} \quad X = \lambda \cdot \left(\frac{1}{\beta_0} + t\right)$$

On a alors :

$$\lambda = \frac{X}{\left(\frac{1}{\beta_0} + t\right)} \quad \text{et} \quad d\lambda = \frac{dX}{\left(\frac{1}{\beta_0} + t\right)}$$

l s'écrit alors :

$$I = \int_0^{+\infty} \left[\frac{X}{\left(\frac{1}{\beta_0} + t\right)} \right]^{\alpha_0+k-1} \cdot e^X \cdot \frac{dX}{\left(\frac{1}{\beta_0} + t\right)}$$

$$I = \frac{1}{\left(\frac{1}{\beta_0} + t\right)^{\alpha_0+k}} \int_0^{+\infty} X^{\alpha_0+k-1} \cdot e^X \cdot dX$$

Or

$$\int_0^{+\infty} X^{\alpha_0+k-1} \cdot e^{-X} \cdot dX = \Gamma(\alpha_0 + k)$$

D'où

$$I = \frac{\Gamma(\alpha_0 + k)}{\left(\frac{1}{\beta_0} + t\right)^{\alpha_0+k}}$$

Ainsi on obtient finalement la densité de probabilité a posteriori :

$$g(\lambda; k, \alpha_0 \beta_0) = \frac{\left(\frac{1}{\beta_0} + t\right)^{\alpha_0+k}}{\Gamma(\alpha_0 + k)} \lambda^{\alpha_0+k-1} \cdot e^{-\lambda\left(\frac{1}{\beta_0}+t\right)}$$

La loi $(\lambda; k, \alpha_0 \beta_0)$ ainsi obtenue est une loi Gamma de paramètres :

$$\left. \begin{aligned} \alpha_{post} &= \alpha_0 + k \\ \beta_{post} &= \frac{1}{\beta_0} + t \end{aligned} \right\}$$

Où α_{post} représente le nombre cumulé de défaillances,

β_{post} représente le temps cumulé de fonctionnement.

Les inférences statistiques déduites à partir de la distribution a posteriori $(\lambda; k, \alpha_0 \beta_0)$ sont les suivantes:

3.3.1 Espérance mathématique:

L'espérance mathématique de la densité de probabilité a posteriori correspond à l'estimateur ponctuel bayésien du taux de défaillance :

$$E(\lambda_p/t; k, \alpha_0, \beta_0) = \hat{\lambda} = \int_{\lambda} \lambda \cdot g(\lambda; k, \alpha_0 \beta_0) d\lambda$$

$$\hat{\lambda}_p = \frac{\alpha_{post}}{\beta_{post}} = \frac{\alpha_0 + k}{\frac{1}{\beta_0} + t} = \frac{\beta_0 \cdot (\alpha_0 + k)}{\beta_0 \cdot t + 1}$$

λ_p est l'estimateur du taux de défaillance a postérieur

3.3.2 La Variance :

$$V(\lambda_p/t; k, \alpha_0, \beta_0) = \frac{\alpha_{post}}{\beta_{post}^2} = \frac{\beta_0^2 \cdot (\alpha_0 + k)}{(\beta_0 \cdot T_c + 1)^2}$$

3.3.3 L'intervalle de crédibilité bayésienne

Un intervalle de crédibilité symétrique bayésien qui contient au moins $(1-\alpha)\%$ de la vraie valeur de λ peut être calculé respectivement à l'aide des relations :

$$\left. \begin{aligned} P[\lambda \leq \lambda_{inf}] &\leq \frac{\alpha}{2} = \int_0^{\lambda_{inf}} g(\lambda; k, \alpha_0 \beta_0) d\lambda, \\ P[\lambda \geq \lambda_{sup}] &\leq \frac{\alpha}{2} = \int_{\lambda_{sup}}^{\infty} g(\lambda; k, \alpha_0 \beta_0) d\lambda \end{aligned} \right\}$$

Ce qui donne en utilisant l'approximation du Chi deux [18-20]:

$$\lambda_{inf} = \frac{x_{\gamma/2}^2 (2\alpha_0)}{2\beta_0} < \lambda < \lambda_{sup} = \frac{x_{1-\gamma/2}^2 (2\alpha_0 + 2)}{2\beta_0}$$

À partir de ces deux expressions, il est possible d'évaluer l'intervalle de crédibilité à 80%, 90%d'un taux de défaillance en fonctionnement.

Exercice :

Etude prévisionnelle: $(1-\gamma=90\%)$ de chance la moyenne de défaillance 0,015 panne/an et au max 0,045 panne/an

Enregistré: 3 pannes sur 1an d'observations sur 220 compteurs

Hypothèse: la fiabilité suit une loi exponentielle

On doit déterminer les paramètres de la loi Gamma, on fait alors une approximation khi-deux:

$$\left\{ \begin{aligned} \lambda_{moy} &= \frac{\alpha_0}{\beta_0} = 0,015 \quad \dots (1) \\ \lambda_{max} &= \frac{x_{1-\gamma/2}^2 (2\alpha_0 + 2)}{2\beta_0} = 0,045 \quad \dots (2) \end{aligned} \right.$$

En remplaçant la première équation dans la deuxième :

$$\begin{cases} \beta_0 = \frac{\alpha_0}{0,015} \\ \frac{x_{1-\gamma/2}^2 (2\alpha_0 + 2)}{2 \frac{\alpha_0}{0,015}} = 0,045 \end{cases}$$

Sachant que $1 - \gamma/2 = 0,95$, Cherchons la valeur de α_0 par approximation (voir la table du Khi-deux dans l'annexe):

On suppose $\alpha_0 = 1$:

$$\frac{x_{0,95}^2 (2+2)}{2 \frac{1}{0,015}} \stackrel{?}{=} 0,045 \rightarrow 0,0075 * 9,49 = 0,071 \neq 0,045$$

$$\alpha_0 = 2:$$

$$\frac{x_{0,95}^2 (4+2)}{2 \frac{2}{0,015}} \stackrel{?}{=} 0,045 \rightarrow 0,0375 * 12,59 = 0,047 \neq 0,045$$

$$\alpha_0 = 3:$$

$$\frac{x_{0,95}^2 (6+2)}{2 \frac{3}{0,015}} \stackrel{?}{=} 0,045 \stackrel{?}{=} 0,045 \rightarrow 0,0025 * 15,51 = 0,039 \neq 0,045$$

De ce fait, la valeur qui se rapproche le plus correspond à:

$$\alpha_0 = 2 \text{ et } \beta_0 = 133$$

Nous pouvons maintenant déduire les paramètres de la loi à posteriori et ainsi le nouveau taux de défaillance :

$$\hat{\lambda}_p = \frac{\alpha_{post}}{\beta_{post}} = \frac{\alpha_0 + k}{\frac{1}{\beta_0} + t} = \frac{\beta_0 \cdot (\alpha_0 + k)}{\beta_0 \cdot t + 1}$$

$$\lambda_p = \frac{133 \left(2 + \frac{3}{220}\right)}{133 * 1 + 1} = 2 \text{ pannes/an}$$

Ce nouveau taux de défaillance élevé nous montre que cet équipement est en phase de vieillissement.

CHAPITRE 7 : Actualisation Bayésienne

1. Introduction

L'actualisation s'applique à des informations homogènes : un retour d'expérience qui intervient comme connaissance a priori et un retour d'expérience plus récent deviens la fonction de vraisemblance.

Ce principe d'actualisation est utilisé dans les banques de données actuelles (EIRéDA, OREDA...)

2. Les conditions et contraintes de l'actualisation

L'actualisation s'applique lorsque l'une ou plusieurs de conditions suivantes se réalisent :

- Lorsque les observations sont nombreuses et en quantité suffisante et égale a chaque fois, ça confirme l'a priori
- Si les résultats d'observations sont différents de l'a priori (des modifications ou vieillissement), alors l'avis de l'expert est requis
- Lorsque les observations récentes ne recueillent aucune défaillance, on fait aussi appel au jugement d'experts pour essayer de donner la loi a priori la plus adéquate

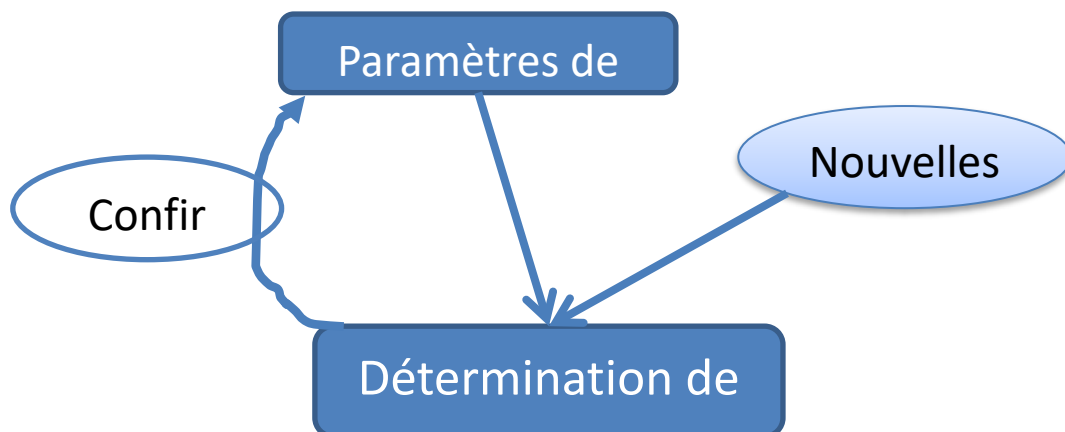


Figure 37 Principe d'actualisation

3. Actualisation des paramètres de fiabilité

Les paramètres de fiabilité établis à partir d'expériences recueillies dans le passé constituent les données apriori : on connaît par exemple le taux de défaillance d'un matériel et son intervalle de confiance.

Les nouvelles observations recueillies plus tard permettent d'actualiser les paramètres de fiabilité à priori, elles représentent les observations de la fonction de vraisemblance.

La distribution jointe à posteriori va confirmer ou infirmer l'étude probabiliste de la fiabilité initiale.

On prendra toujours comme exemple taux de défaillance, l'actualisation des autres paramètres de fiabilité suit les mêmes étapes.

On supposera aussi que les défaillances suivent une loi exponentielle. Ceci concerne surtout le matériel assujettis à une maintenance préventive, la distribution a priori est une loi Gamma (conjuguais à une loi exponentielle).

Comme pour la modélisation de la connaissance (loi à priori dans le chapitre précédent) on cherche à définir les paramètres à priori α et β pour pouvoir déterminer les paramètres à postériori et qui sera par la suite une nouvelle connaissance à priori et de ce fait, une actualisation Bayésienne de la connaissance.

4. La méthode des moments

Nous pouvons procéder, comme dans le chapitre précédent, par approximation du Khi-deux, mais dans [18] une autre méthode est proposée, c'est la méthode du moment.

Cette méthode consiste à identifier les deux premiers moments de la loi apriori Gamma avec les estimateur moyenne et variance obtenus par l'expertise ou le retour d'expérience :

$$\begin{cases} E[\lambda] = \alpha \cdot \beta \\ Var[\lambda] = \alpha \cdot \beta^2 = \sigma^2 \end{cases}$$

Sachant que : $2.1,64 \cdot \sigma = (\lambda_{sup} - \lambda_{inf})_{90\%}$

Alors on a le système d'équation :

$$\begin{cases} E[\lambda] = \alpha \cdot \beta \\ 2.1,64 \cdot \sigma = (\lambda_{sup} - \lambda_{inf}) \end{cases}$$

Ceci, bien entendu, permettra de déterminer le taux de défaillance a posteriori :

$$\hat{\lambda} = E[\lambda|k, T] = \frac{\beta(\alpha + k)}{\beta T + 1}$$

Exemple :

Soit :

Le Taux de défaillance moyen $93 \cdot 10^{-6}$ L'intervalle de confiance est $[56,3 \cdot 10^{-6} ; 146 \cdot 10^{-6}]$ à 90%.

Les nouvelles observations :

K = 20 sur $2,71 \cdot 10^5$ heures

$$2.1,64 \cdot \sigma = (\lambda_{sup} - \lambda_{inf})_{90\%} \rightarrow \sigma = \frac{(\lambda_{sup} - \lambda_{inf})_{90\%}}{2.1,64} = \frac{146 \cdot 10^{-6} - 56,3 \cdot 10^{-6}}{3,28} = 27,3510^{-6}$$

Et du système d'équation :

$$\left\{ \begin{array}{l} E[\lambda] = \alpha \cdot \beta \\ \text{Var}[\lambda] = \alpha \cdot \beta^2 = \sigma^2 \end{array} \right. \rightarrow \left\{ \begin{array}{l} E[\lambda] = \alpha \cdot \beta \\ \text{Var}[\lambda] = \alpha = \frac{\sigma^2}{\beta^2} \end{array} \right. \rightarrow \left\{ \begin{array}{l} E[\lambda] = \frac{\sigma^2}{\beta} \\ \text{Var}[\lambda] = \alpha = \frac{\sigma^2}{\beta^2} \end{array} \right.$$

$$E[\lambda] = \frac{(27,35 \cdot 10^{-6})^2}{\beta} = 93 \cdot 10^{-6} \rightarrow \beta = 8,01 \cdot 10^{-6}$$

$$\alpha = \frac{\sigma^2}{\beta^2} = \frac{(27,35 \cdot 10^{-6})^2}{(8,01 \cdot 10^{-6})^2} = 12$$

Ceci permet de déterminer le taux de défaillance a posteriori :

$$\hat{\lambda} = E[\lambda|k, T] = \frac{\beta(\alpha+k)}{\beta T+1} = \frac{8,01 \cdot 10^{-6} (12+20)}{8,01 \cdot 10^{-6} * 2,71 \cdot 10^5 + 1} =$$

Références (Par ordre d'apparition)

- [1] INERIS-DRA-2006-P46055-CL47569 : $\Omega 7$: Méthodes d'analyse des risques générés par une installation industrielle
- [2] Pagetti.C, (2012) Module de sûreté de fonctionnement – ENSEEIHT
- [3] Alain Villemeur, Sûreté de fonctionnement des systèmes industriels, Paris, Eyrolles, coll. « Collection de la direction des études et recherches d'Électricité de France », juillet 1988 (ISSN 0399-4198) page 744
- [4] Gilles ZWINGELSTEIN, Sûreté de fonctionnement des systèmes industriels complexes, Techniques de l'Ingénieur, traité Informatique industrielle
- [5] David, P. (2011), cours : Management des Risques Industriels Déploiement de la Sûreté de Fonctionnement: Notions, Méthodes, Cycle de Vie, Génie Industrielle Grenoble
- [6] Apostolakis A (1990), The Concept of Probability in Safety Assessment of Technological Systems, Science, 250 : 1359-1364
- [7] Winkler RL (1996), Uncertainty in probabilistic risk assessment, Reliability Engineering & System Safety, 54 : 127-132
- [8] Sallak M (2008), Evaluation de paramètres de sûreté de fonctionnement en présence d'incertitudes et aide à la conception : application aux Systèmes Instrumentés de Sécurité, Doctorat de l'Institut National Polytechnique de Lorraine
- [9] Tebbi O (2005), Estimation des lois de fiabilité en Mécanique par les essais accélérés, ECOLE DOCTORALE D'ANGERS
- [10] Robert, C. (1992). L'analyse statistique bayésienne. Edition ECONOMICA.
- [11] Ibrahim, J., Chen, M., and Sinha, D. (2001). Bayesian survival analysis. Ed Springer.
- [12] Ringler, J. (1979). Utilisation des techniques bayésiennes dans le cas des systèmes complexes. Revue de Statistique Appliquée, vol XXVII n°2
- [13] Lannoy, A. and Procaccia, H. (2001). L'utilisation du jugement d'expert en sûreté de fonctionnement. Ed. TEC&DOC.
- [14] Pearl, Judea, and Stuart Russell. "Bayesian Networks." UCLA Cognitive Systems Laboratory, November 2000. http://bayes.cs.ucla.edu/csl_papers.html.
- [15] Anass BOUCHIBA., (Evaluation De Dysfonctionnement D'un Système Par Approche Bayésienne : Cas Du Système Ferroviaire), Thèse de doctorat, Institut des Sciences et Techniques de l'Ingénieur d'Angers, 2013.

[16] www.bayesia.com, (Knowledge Modeling Quick Star Guide), 08/01/2015

[17] H. Procaccia. "Introduction à l'analyse probabiliste des risques industriels". Editions Tec & Doc, ISBN : 978-2-7430-1100-7, Lavoisier 2009.

[18] PROCACCIA Henri, SUHNER Marie-Christine, "Démarche bayésienne et applications à la sûreté de fonctionnement », Éditions Lavoisier 2003.