وزارة التعليم العالي والبحث العلمي

Université Badji Mokhtar
Annaba

Badji Mokhtar University -
Annaba

جامعة باجي مختار
عنابة

Faculté des Sciences
Département de Mathématiques

THESE

Présenté en vue de l'obtention du diplôme de
Doctorat 3ème cycle en Mathématiques
Option : Probabilité et Statistiques

o

Estimation et Prédiction
Dans les problèmes de durée de survie

Par:
MERADJI Asma

Devant le jury

Présidente :	**Seddik-Ameur Nacira**	**Prof (U.B.M.Annaba)**
Rapporteur :	**Chadli Assia**	**MCA (U.B.M.Annaba)**
Examinatrice :	**Messaci Fatiha**	**Prof (U.Constantine)**
Examinatrice :	**Nemouchi Nahima**	**Prof (U.Constantine)**

Année: 2015

Table des matières

0.1 Abstract

In this thesis, we focus on the study of statistical inference in two survival models, namely the exponential model and Rayleigh model. This study deals firstly with a parameter estimation problems, function reliability and failure rate for both mentioned models and also the problem of prediction order statistics and functions of the latter for the exponential model. Indeed, the prediction order statistic functions allow the deduction of prediction lifetime of some complex systems.

The approach used to the exponential model is a Bayesian approach, combined with a quadratic loss function. Three types of data were considered ; these are the type II censored data, attribute data type for which the time of observations is fixed in advance and grouped data. A simulation study was conducted to check all the results.

For the Rayleigh model, a panel of loss function, and in particular the asymmetric loss functions were used to estimate the parameter of the function of reliability and failure rate. The data comes from an experimental design right censored. In the first part, we considered the case of a prior law uninformative ; then we studied the case of a law natural conjugate prior. The Gibbs sampling algorithm was used to perform simulations to calculate the different estimators.

.

0.2 Résumé

Dans cette thèse, nous nous attachons à l'étude de l'inférence statistique dans deux modèles de survie, à savoir, le modèle exponentiel et le modèle de Rayleigh. Cette étude traite d'une part des problèmes d'estimation des paramètres, de la fonction de fiabilité et du taux de panne pour les deux modèles cités çi-dessus et d'autre part au problème de la prédiction de statistiques d'ordre et de fonctions de celles-ci pour le modèle exponentiel. En effet, la prédiction des fonctions de statistique d'ordre permet de déduire la prédiction de la durée de vie de quelques systèmes complexes.

L'approche utilisée pour le modèle exponentiel est une approche Bayesienne, assiciée à une fonction de perte quadratique. Trois types de données ont été considérées ; il s'agit des données censurées de type II, des données de type attribut pour lesquelles le temps d'observations est fixé au préalable et des données groupées. Une étude par simulation a été menée pour vérifier tous les résultats obtenus.

Pour le modèle de Rayleigh, un panel de fonctions de perte, et en particulier les fonction de perte asymétriques ont été utilisées pour l'estimation du paramètre, de la fonction de fiabilité et du taux de panne. Les données proviennent d'un plan d'expérience censuré à droite. Dans une première partie, nous avons considéré le cas d'une loi a priori non informative ; puis nous avons étudié le cas d'une loi a priori conjuguée naturelle. L'algorithme d'échantillonnage de Gibbs a été utilisé pour réaliser des simulations afin de calculer les différents estimateurs.

0.3 Remerciements

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

En premier lieu, je tiens à remercier ma directrice de thèse, Madame Assia CHADLI, pour la confiance qu'elle m'a accordée en acceptant d'encadrer ce travail doctoral, pour ses multiples conseils et pour toutes les heures qu'elle a consacrées à diriger cette recherche. J'aimerais également lui dire à quel point j'ai apprécié sa grande disponibilité et son respect sans faille des délais serrés de relecture des documents que je lui ai adressés. Enfin, j'ai été extrêmement sensible à ses qualités humaines d'écoute et de compréhension tout au long de ce travail doctoral.

Je suis grandement reconnaissante à Madame Nacira SEDDIK AMEUR, Professeur à l'université Badji Mokhtar de Annaba pour avoir accepté de présider mon jury de thèse.

Les examinatrices qui ont accepté de siéger dans le jury de cette thèse doivent aussi trouver ici l'expression de ma reconnaissance, soit, Mme Fatiha MESSACI, et Mme Nahima NEMOUCHI ; Professeurs à l'université Mentouri de Constantine et Mme Natalia.DJELLAB Professeur à l'université Badji Mokhtar de Annaba.

Mes remerciements vont particulièrement à mon père, qui m'a élevé avec ce grand goût de dépassement. J'adresse des remerciements de même ordre à ma mère, qui m'a constamment encouragée et soutenue tout au long de ces années. Je ne saurai passer sous silence l'apport inestimable de mon mari ,ma fille la rose de ma vie 'Serine'et des autres membres de ma famille (frères, sœurs).

0.4 Introduction

Pour évaluer la fiabilité de leurs produits, les entreprises (s'il s'agit de **fiabilité**) ou les épidémiologistes (s'il s'agit de **durée de survie**) ; les deux termes signifiant la même entité mathématique doivent mettre en œuvre des méthodes d'évaluation appropriées pour l'estimation des paramètres et des différentes caractéristiques de fiabilité. Ces méthodes d'évaluation sont basées sur des données provenant de :

1) Banques de données de fiabilité issues de l'historique des composants étudiés, et éventuellement de la simulation de leur fonctionnement. Ces données peuvent être complètes ou censurées ; pour des raisons évidentes de coût, les données complètes sont rarement utilisées.

2) Données issues d'avis d'experts.

Le problème de la prédiction de statistiques d'ordre et de fonctions de celles-çi et ses applications en fiabilité a fait l'objet de plusieurs articles. Plusieurs auteurs se sont intéressés à ce problème en utilisant une approche classique ; Dunsmore (1974) ; Bancroft G.A. (1976) ; Prezzi (1996); Pan JN. (1997). L'approche Bayesienne a été également utilisée avec des lois à priori conjuguées naturelles ou non informative ; des données complètes ou censurées et différentes lois. Bratcher (1971) a utilisé des données complètes ; Dunsmore (1974) a étudié d'une manière générale les tests de survie ; Evans et Nigm (1980) ont étudié la loi exponentielle tronquée et la loi exponentielle bivariée avec des données censurées de type II ; Arnold BC. (1989) a considéré la loi de Pareto ; Erkanly (1998) et Van-batenberg (1989) ont appliqué les résultats de la prédiction Bayesienne à la fiabilité. Basu et al (2003) ont étudié un modèle soumis à des risques de défaillances concourants, Jozani, M.J (2012) a comparé à l'aide du critère de Pittman les estimateurs obtenus avec différentes fonctions de perte.

Cette thèse présente différentes approches qui consistent à estimer les différents paramètres d'un modèle de survie. Outre, les paramètres, on s'est interéssé en particulier à l'estimation de la fonction de fiabilité et à l'estimation du taux de panne. Les modèles étudiés en détail sont les **modèles exponentiels** et les **modèles de Rayleigh**. L'approche utilisée pour ce problème d'inférences statistique est une approche bayesienne. Le choix de la loi a priori est souvent perçu comme une difficulté majeure de l'approche bayesienne en ce que l'interprétation de l'information a priori disponible est rarement assez précise pour conduire à la détermination d'une seule et unique loi. On utilisera dans cette thèse deux types de loi a priori, qui sont les loi a priori *conjuguées naturelles* et les lois a priori *non informatives.*

La fonction de perte illustre la relation entre la baisse de performance et les coûts que celle-ci entraîne pour le client et la société en général. On considère en effet que tout écart par rapport à la valeur optimale engendre des pertes, car le produit n'offre plus les performances optimales. Cette fonction peut aussi être utilisée pour définir les spécifications et les tolérances en fonction de l'analyse des coûts. Les estimateurs bayesiens sont directement liés au choix de la fonction de perte, qui est généralement choisie de type quadratique, cependant plusieurs fonctions de pertes ont été développées ces dernières années, en particulier les fonctions de perte asymétriques LINEX. Dans cette thèse, nous utiliserons un panel de fonctions de perte pour l'estimation des paramètres, de la fonction de fiabilité et du taux de panne dans un modèle de Rayleigh.

Organisation du manuscrit

Ce manuscrit de thèse est organisé en trois chapitres de part plus ou moins égales.

Chapitre 1 : Ce chapitre est consacré aux rappels des outils mathématiques nécessaires à l'aboutissement de ce travail, on y définit les notions :

1) De fiabilité et ses principales caractéristiques

2) D'estimation et de prédiction à l'aide d'une approche classique du maximum de vraisemblance, puis à l'aide d'une approche Bayesienne

3) Les différents plans d'expériences utilisés dans cette thèse

4) Un panel de fonctions de perte

5) Statistiques d'ordre.

Chapitre 2 : Dans ce chapitre, on s'interesse au modèle exponentiel ; on rappelle les résultats obtenus par Dunsmore (1997) et Evans et Nigm (1998) ,dans le cas de données censurées de type II ; les résultats obtenus par Chadli (2004) dans le cas des données de type attribut et on étudiera le cas des données groupées. (ce travail a fait l'objet d'une publication dans la revue synthèse).

Chapitre 3 : Ce chapitre est consacré au modèle de Rayleigh. On s'est intéressé à l'estimation du paramètre, de la fonction de fiabilité et du taux de panne dans le cas de données censurées de type II pour une loi a priori non informative puis on consiéré le cas d'une loi a priori conjuguée naturelle. La thèse est achevée par une **conclusion** et quelques **perspectives** de développement sur l'utilisation des données progressivement censurées.

Chapitre 1

Notions de base

1.1 Caractéristiques de fiabilité

Dans ce qui suit, nous considérons un individu susceptible de subir une fois et une seule un certain type dévénement (décès ou avarie par exemple). L'observation de la survenue - donnée non censurée - on non - donnée censurée- de cet événement chez l'individu constitue la donnée basale pour une modélisation de la survie.

Nous donnons ci-dessous les définitions des principaux outils utilisés en analyse de la survie.

Definition 1 *La " durée de vie" d'un individu est une variable aléatoire (v.a) X positive et continue. Sa fonction de répartition est la probabilité que l'événement se produise entre 0 et x.*

$$F(x) = \mathbb{P}(X \leq x) \tag{1.1}$$

Definition 2 *La fonction de survie ou fonction de fiabilité est définie par*

$$\begin{aligned} S(x) &= 1 - F(x) \\ &= \mathbb{P}(X > x). \end{aligned} \tag{1.2}$$

Definition 3 *La fonction de risque instantané ou taux de hasard est la fonction*

$$h(x) = \frac{f(x)}{S(x)} \tag{1.3}$$

où f est la densité de probabilité de X.

Definition 4 *La fonction de risque cumulé est donnée partions de base*

où f est la densité de probabilité de X.

Definition 5 *La fonction de risque cumulé est donnée par*

$$H(x) = \int_0^x h(u)du \tag{1.4}$$

Nous pouvons maintenant énoncer les deux résultats suivants :

Proposition 6 *La définition de la distribution de probabilité de X repose sur l'une des quatre données suivantes, qui sont équivalententes :*$S(x), f(x), h(x)$ *et* $H(x)$.

Proposition 7 *Nous avons*

$$\begin{aligned} S(x) &= \mathbb{P}(X > x). \qquad (1.5)\\ &= \underset{]0,x]}{\mathcal{P}}\left[1 - h(s)ds\right] \\ &= \exp\left(-\int_0^x h(s)ds\right) \\ &= \exp\left(-H(x)\right). \end{aligned}$$

$\underset{]0,x]}{\mathcal{P}}$ *et désigne le produit infini (ou produit intégrale).*

1.2 Modèles usuels en fiabilité

Nous allons présenter dans cette section quelques lois qui peuvent être utilisées dans le cadre d'une étude sur des données de survie. Bien que n'importe quelle distribution d'une variable aléatoire continue non-négative puisse être utilisée, nous ne parlons ici que des lois les plus usitées.

1.2.1 Modèle exponentiel

Il s'agit du modèle le plus simple : on postule que le risque instantané est une constante

$$h(t) = \lambda, \quad t \geq 0, \lambda > 0 \tag{1.6}$$

Compte tenu de la relation fondamentale entre la fonction de risque et la fonction de survie

$$S(t) = \exp\left[-\int_0^t \lambda du\right] = e^{-\lambda t} \tag{1.7}$$

Soit encore :

$$log(S(t)) = -\lambda t$$

En ce qui concerne la fonction de densité des temps de survie on a :

$$f(t) = \frac{ds}{dt} = \lambda e^{-\lambda t} \tag{1.8}$$

Expression où on reconnaît la fonction de densité d'une variable exponentielle de parametre λ.En d'autre termes, une fonction de risque constante implique une distribution exponentielle des durées d'évènement.

1.2.2 Modèle de Weibull

La distribution de Weibull est egalement une généralisation de l'exponentielle. Elle est caractérisée par deux paramètres,$\beta > 0, \eta > 0$ qui sont les paramètres de forme et d'échelle respectivement. La fonction de densité d'une telle loi est donnée par

$$f(t, \beta, \eta) = \frac{\beta}{\eta^\beta} t^{\beta-1} \exp[-(\frac{t}{\eta})^\beta], \quad t > 0. \tag{1.9}$$

Ainssi nous remarquons que si $\beta = 1$,nous obtenons la loi exponentielle. La fonction de survie est $S(t) = \exp[-(\frac{t}{\eta})^\beta]$ et le taux de panne vaut $h(t) = \frac{\beta}{\eta}\left(\frac{t}{\eta}\right)^{\beta-1}$, qui est donc de l'ordre de $t^{\beta-1}$.

La fonction de risque est monotone croissante si $\beta > 1$, monotone décroissante si $\beta < 1$ et constante pour $\beta = 1$, c'est pourquoi ce paramètre est appelé paramètre de forme. Comme η est un paramètre d'échelle, différentes valeurs de η changent seulement l'échelle sur l'axe horizontal, et non pas la forme de base du graphe. Le modèle est assez flexible, et on a montré qu'il constitue une bonne description de plusieurs types de données de survie. Le fait que les fonctions de densité, de survie et de risque aient une forme relativement simple explique également la popularité du modèle.

Soit T une variable aléatoire suivant une loi de Weibull, sa moyenne et sa variance sont données par les formules suivantes :

$$\begin{aligned} E(T) &= \eta\Gamma(1 + \frac{1}{\beta}) \\ Var(T) &= \eta^2\Gamma(1 + \frac{2}{\beta}) - E(T)^2 \end{aligned}$$

Où $\Gamma(x) = \int_0^{+\infty} t^{x-1}e^{-t}dt$ est la fonction gamma.

L'on parle de loi de Weibull décalée lorsque l'on considère que le taux est constant et égale à zéro jusqu'à un instant t_0.Dans ce cas, la fiabilité et le taux de hasard s'écrivent de la façon suivante :

$$\begin{aligned} S(t) &= \exp[-(\frac{t-t_0}{\eta})^{\beta}] \qquad (1.10)\\ h(t) &= \frac{\beta}{\eta}\left(\frac{t-t_0}{\eta}\right)^{\beta-1} \end{aligned}$$

1.2.3 Modèle Log-Normale

La durée de vie T a une distribution log-normale si $Y = \log(T)$ est une distribution normale.

Ainsi, si Y est une gaussienne d'espérance μ_Y et de variance σ_Y^2 et donc de densité

$$\phi(y) = \frac{1}{\sigma_Y\sqrt{2\pi}} \exp\left[-\frac{1}{2}\left(\frac{y-\mu_Y}{\sigma_Y}\right)^2\right], \quad -\infty \ < y < \infty \qquad (1.11)$$

alors,

$$f(t) = \frac{1}{t\sigma_Y\sqrt{2\pi}} \exp\left[-\frac{1}{2}\left(\frac{\log(t)-\mu_Y}{\sigma_Y}\right)^2\right],\ t > 0. \qquad (1.12)$$

où μ est le paramètre d'échelle et σ est le paramètre de forme. Contrairement à la loi normale, les paramètres ne donnent pas la moyenne et la variance de la loi.

et la fonction de survie d'une variable suivantune loi log-normale est donnée par

$$S(t) = 1 - \Phi\left(\frac{\log(t)-\mu_Y}{\sigma_Y}\right) \qquad (1.13)$$

Où $\Phi(.)$ est la fonction de répartition d'une gaussienne centrée-réduite,

$$\Phi(x) = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}} e^{-v^2} dv \qquad (1.14)$$

Le taux de panne est de la forme

$$h(t) = \frac{\frac{1}{\sigma_Y\sqrt{2\pi}} \exp\left[-\frac{1}{2}\left(\frac{y-\mu_Y}{\sigma_Y}\right)^2\right]}{1-\Phi\left(\frac{\log(t)-\mu_Y}{\sigma_Y}\right)} \qquad (1.15)$$

Pour évaluer $S(t)$ et donc aussi $h(t)$ il est nécessaire d'évaluer numériquement l'intégrale ce qui affecte les temps de calcul.

Avec la distribution log-normale, la fonction de risque $h(t)$ est croissante puis décroissante avec $h(0) = 0$ et $\lim_{t \to \infty} h(t) = 0$.

Comme la fonction de risque décroît pour de grandes valeurs de t, la distribution ne paraît pas plausible comme modèle de vie dans la plupart des situations. Malgré cela, ce modèle peut être intéressant lorsque de très grandes valeurs de t ne sont pas d'un intérêt particulier.

1.2.4 Modèle Gamma

La loi gamma comporte deux paramètres, $\lambda > 0$ et $\gamma > 0$ Le premier est appelé paramètre d'échelle alors que γ est le paramètre de forme. La fonction de densité de cette loi est donnée par

$$f(t; \lambda, \gamma) = \frac{\lambda}{\Gamma(\gamma)} (\lambda t)^{\gamma-1} e^{-\lambda t}, \quad t > 0. \tag{1.16}$$

où $\Gamma(\gamma) = \int_0^\infty x^{\gamma-1} e^{-x} dx$ est la fonction Gamma. La fonction de survie s'exprime comme

$$S(t) = \int_0^\infty \frac{\lambda}{\Gamma(\gamma)} (\lambda x)^{\gamma-1} e^{-\lambda x} dx \tag{1.17}$$

En choisissant le paramètre γ entier, nous obtenons la distribution dite de Erlang. Pour celle-ci, nous obtenons comme taux de panne

$$h(t) = \frac{\lambda (\lambda t)^{\gamma-1}}{(\gamma - 1)! \sum_{k=0}^{\gamma-1} (\lambda t)^k} \tag{1.18}$$

Il est facile de montrer que la fonction de risque $h(t)$ monotone croissante pour $\gamma > 1$.avec $\lim_{t \to \infty} h(t) = \lambda$.pour $0 < \gamma < 1, h(t)$ est monotone décroissante avec $\lim_{t \to 0^+} h(t) = \infty$ et $\lim_{t \to \infty} h(t) = \lambda$, Si l'on choisit $\gamma = 1$, nous obtenons une loi exponentielle de paramètre λ,ainsi, la loi exponentielle n'est qu'un cas particulier de la loi gamma.

La distribution gamma apparaît aussi dans des cas où la distribution exponentielle est utilisée, car la somme de variable aléatoires exponentielles indépendantes et identiquement distribuées suit une distribution gamma.

Plus précisement, si $T_{1,} T_{2,} ..., T_n$ sont des variables aléatoires indépendantes suivant une loi exponentielle de paramètre λ, alors $\sum_{i=1}^n T_i$ suit une distribution gamma de paramètres λ et $\gamma = n$.

Le logarithme de la vraisemblance d'un échantillon issu d'une loi gamma est donné par

$$l(\lambda, \gamma) = n\lambda \ln \lambda - n \ln \Gamma(\gamma) + (\gamma - 1) \sum_{i=1}^n \ln t_i - n\lambda \bar{t}$$

En dérivant par rapport à λ et en appelant $\widehat{\gamma}$ l'estimateur du maximum de vraisemblance pour γ,nous obtenons

$$\widehat{\lambda} = \frac{\widehat{\gamma}}{\overline{t}}$$

Par contre, le calcule exacte de $\widehat{\gamma}$ n'est pas possible, ainssi, nous pouvons seulement exprimer un estimateur en fonction de l'autre.

1.3 Méthodes d'estimations paramétriques

1.3.1 Définition informelle d'un estimateur et d'une estimation

Dans le cadre défini ci dessus, soit à estimer une caractéristique c de la variable aléatoire X sur la base de la réalisation $(x_1, x_2, ..., x_n)$ du n-échantillon $(X_1, X_2, ..., X_n)$. On appellera estimateur toute statistique (donc toute fonction de $X_1, X_2, ..., X_n$) dont la réalisation aprés expériance est envisagée comme estimation de c. Un estimateur se définit donc dans l'intention de fournir une estimation.

Insistons sur le fait qu'un estimateur est une variable aléatoire , alors qu'une estimation est une valeur numérique prise par l'estimateur suite à la réalisation du n-échantillon . Si un estimateur est déterminé par une fonction $h(X_1, X_2, ..., X_n)$, l'estimation correspondante est évidemment $h(x_1, x_2, ..., x_n)$. Soit, par exemple, à estimer la moyenne $E(X)$ de la loi de X, un estimateur qui semble a priori naturel est la moyenne empirique $\overline{X}$ qui produit une estimation $\overline{x}$, moyenne descriptive de la série des valeurs observées.

1.3.2 Cadre de l'estimation paramétrique

En estimation paramétrique, la loi X est réputée appartenir à une famille de lois, qui peut être décrite par une forme fonctionnelle qui peut être sa fonction de répartition $F(x;\theta)$ ou sa fonction de densité $f(x;\theta)$.

L'ensemble des valeurs possibles pour θ , appelé espace paramétrique, sera noté Θ ,lequel est inclus dans $\mathbb{R}^k$ où k est la dimension du paramètre θ.Le plus souvent la famille paramétrique à laquelle la loi de X est réputée appartenir sera décrite par la famille de densités de probabilité (respectivement de fonctions de probabilité) $\{f(x;\theta);\theta \in \Theta\}$. Ces fonctions sont implicitement définies pour tout $x \in \mathbb{R}$. Rappelons ici qu'on appelle support de $f(x;\theta)$ (ou support de la loi) l'ensemble des valeurs de x telles que $f(x;\theta) > 0$. En termes courants, on parlera de l'ensemble des réalisations (ou valeurs) possibles.

Pour clore cette introduction mettons en evidence le fait que, dans le cadre des échantillons aléatoires, la loi conjointe du n-échantillon $(X_1, X_2, ..., X_n)$

peut être définie par la fonction de densité (respectivement la fonction de probabilité) conjointe :

$$f(x_1, x_2, ..., x_n; \theta) = f(x_1; \theta) f(x_2; \theta) ... f(x_n; \theta) = \prod_{i=1}^{n} f(x_i; \theta)$$

où θ est le paramètre inconnu dans Θ (par commodité, nous désignons ici, et ferons de même éventuellement plus loin, la densité conjointe par la même lettre f que la densité de la loi mère). Dans le cas discret cette expression est la probabilité $P_0 (X_1 = x_1, X_2 = x_2, ..., X_n = x_n)$.

Pour établir certains résultats nous serons amenés à poser des conditions sur cette densité (ou fonction de probabilité) conjointe.Ces conditions ne seront pertinentes qu'aux points $(x_1, x_2, ..., x_n)$ de $\mathbb{R}^n$ correspondant à des va-

leurs possibles, en d'autres termes uniquemnet sur le support de la loi conjointe.Ce support est evidemment l'ensemble produit n fois du support de $f(x; \theta)$.Pour la famille paramétrique dans son ensemble, il faudra prendre en compte l'union des supports de tous les membres lorsque θ décrit Θ.Pour la famille des lois $U[0, \theta]$, avec $\theta > 0$,par exemple, cette union est $\mathbb{R}^+$. Donc, par la suite, les

conditions imposées aux densités ou fonctions de probabilités seront implicitement restreintes à leurs supports.

mettons en évidence le fait que, dans le cadre des échantillons aléatoires, la loi conjointe du n-échantillon $(X_1, X_2, ..., X_n)$

Nous abordons maintenant trois méthodes générales

1.3.3 l'estimation par la méthode du maximum de vraisemblance :

Nous commençons par la méthode du maximum de vraisemblance qui est la plus universelle (y compris pour des modèles complexes) pour deux raisons :

1- Elle est facile à mettre en oeuvre, se ramenant à un problème classique de résolution numérique.

2- D'un point de vue pratique, pour un échantillon suffisament grand (disons $n > 30$ pour fixer les idées), elle fournit des estimateurs de très bonne qualité.

Definition 8 *Soit un échantillon aléatoire $(X_1, X_2, ..., Xn)$ dont la loi mère appartient à une famille paramétrique de densités (ou fonctions de probabilité) $\{f(x; \theta) \ , \theta \in \Theta\}$*

où $\Theta \subseteq \Re^k$.On appelle fonction de vraisemblance de θ pour une réalisation donnée $(x_1, x_2, ..., x_n)$ de l'échantillon, la fonction de θ :

$$L(\theta; x_1, x_2, ..., x_n) = \prod_{i=1}^{n} f(x_i, \theta) \tag{1.19}$$

L'éxpression dé la fonction dé vraisemblancé ést donc la méme que celle de la densité (ou fonction dé probabilité) conjointe mais lé point de vue ést différent. Ici lés valeurs $x_1, x_2, ..., x_n$ sont fixées (ce seront les valeurs éffectivement observées) et on s'intéresse à la façon dont varie la valeur de la densité (ou fonction de probabilité) associée à une sérié d'observations donnée suivant les différentes valeurs de θ. Dans le cas discret il s'agit directement de la probabilité $P_\theta(X_1 = x_1, X_2 = x_2, ..., Xn = x_n)$.S'il n'y a pas d'ambiguité possible, on notera la fonction de vraisemblance simplément $L(\theta)$. On dira que la valeur $\theta_1$de θ est plus vraisemblable que la valeur θ_2 si $L(\theta_1) > L(\theta_2)$. En ce sens il dévient naturel de choisir pour θ la valeur la plus vraisemblable, disons $\hat{\theta}^{MV}$, c'est-à-dire telle que la loi $f\left(x, \hat{\theta}^{MV}\right)$ corréspondante confère la plus forte probabilité (ou densité de probabilité) aux observations rélévéés.

Definition 9 *On appelle estimation du maximum de vraisemblance une valeur $\hat{\theta}^{MV}$, s'il en existe une, telle que :*

$$L\left(\hat{\theta}^{MV}\right) = \sup_{\theta \in \Theta} L(\theta)$$

Une telle solution est fonction de $(x_1, x_2, ..., x_n)$, soit $\hat{\theta}^{MV} = h(x_1, x_2, ..., x_n)$. Cette fonction h induit la statistique (notée abusivement, mais commodément, avec le même symbole que l'estimation) $\hat{\theta}^{MV} = h(X_1, X_2, ..., Xn)$ appelée estimateur du maximum de vraisemblance (EMV).

Cette définition appelle quelques remarques :

- $\hat{\theta}^{MV}$ est une fonction dé $\Re^n$ dans $\Re^k$, associant a tout échantillon particulier une valeur particulière de θ ;

– généralement l'EMV existe et il est unique, i.e. quel que soit $(x_1, x_2, ..., x_n)$ il y a un et un seul maximum pour $L(\theta)$.

-la définition de l'EMV s'étend a des variables aléatoires non i.i.d. car elle ne repose que sur la notion de densité (fonction de probabilité) conjointe. Ellé s'étend même dans un cadre non paramétrique ;

– une fois la réalisation $(x_1, x_2, ..., x_n)$ observée, l'estimation est facilement obtenue, y compris pour des situations complexes. Il suffit d'utiliser un algorithme de maximisation numerique comme on en trouve dans tous les logiciels mathématiques.

Quand les densités (fonctions de probabilité) conjointes sont des produits de fonctions puissances et exponentielles, ce qui est le cas la plupart du temps, on a plutôt intérêt à maximiser $\ln L(\theta)$, appelée log-vraisemblance, ce qui est équivalent puisque la fonction logarithmique est strictement croissante. Dans les cas «réguliers» où $L(\theta)$ est continûment dérivable et le support pour la famille de lois considérée est indépendant de θ, l'estimation par le maximum de vraisemblance (MV) vérifie (pour $\Theta \subseteq \Re$) :

$$\begin{aligned}\frac{\partial}{\partial\theta}\ln L(\theta) &= 0\\ \frac{\partial^2}{\partial\theta^2}\ln L(\theta) &\prec 0\end{aligned}$$

Où

$$\frac{\partial}{\partial\theta}\ln\left[\prod_{i=1}^{n}f(x_i,\theta)\right]=0$$

0ù

$$\sum_{i=1}^{n}\frac{\partial}{\partial\theta}\ln f(x_i,\theta)=0.$$

Cette dernière égalité s'appelle l'équation de vraisemblance. Dans le cas où θ possède k dimensions $(\theta_1,\theta_2,...,\theta_k)$, on résout un système de k équations obtenues en dérivant par rapport à chacune des composantes. Mathématiquement, le fait d'être solution de l'équation (ou du système d'équations) de vraisemblance n'est pas une condition suffisante pour être un maximum. Toutefois étant donne que $L(\theta)$ admet une borne supérieure en tant que probabilité (cas discret) mais aussi, généralement, en tant que densité de probabilité (cas continu), et qu'elle est le plus souvent concave, l'équation admettra une solution unique qui sera alors nécessairement un maximum. Dans les exemples, pour alléger l'exposé, nous n'examinerons pas dans le détail si la solution de l'équation (ou du système d'équations) correspond effectivement à un maximum.

1.4 L'approche Bayesienne

1.4.1 Estimation Bayesienne

Nous abordons ici l'approche bayesienne qui relève d'une philosophie particulière de la statistique. D'une façon générale on qualifie ainsi toute approche qui *confère a tout paramètre inconnu un statut de variable aléatoire* en stipulant pour celui-ci une distribution sur Θ appelée *loi a priori*. Cette loi peut résulter de la connaissance que l'on peut avoir acquise antérieurement sur le phénomène ou être un simple artifice permettant de mener à bien les calculs. En général on tendra à utiliser une *loi a priori* à laquelle le résultat final sera relativement peu sensible (on définit notamment des lois a priori dites «non informatives»). L'espace paramétrique etant généralement continu, définissons cette loi par une densité, notée$\pi(\theta)$. Pour simplifier on supposera le paramètre de dimension 1, mais l'extension à une dimension quelconque ne présente pas de difficultés.

Dans ce cadre,$f(x;\theta)$ doit être considérée maintenant comme une densité (ou fonction de probabilite) *conditionnelle* pour la v.a. X étudiée, étant donné une valeur fixée du paramètre θ *(il serait donc approprié de l'écrire* $f(x|\theta)$). En suivant la formule de Bayes qui permet de passer de la loi de probabilité d'un événement A sachant B à la probabilite de B sachant A selon :

$$P(B/A) = \frac{P(A/B)\,P(B)}{P(A)}$$

on définit la *loi* a posteriori de θ, c'est-à-dire aprés avoir pris connaissance des réalisations $(x_1, x_2, ..., x_n)$ de l'échantillon $(X_1, X_2..., X_n)$. Ci-après le vecteur des réalisations sera noté x et l'echantillon sera noté X. Par transcription de la formule de Bayes la densité a posteriori est :

$$\pi_{\theta/X=x}(\theta) = \frac{f(x;\theta)\,\pi(\theta)}{\int_\Theta f(x;\theta)\,\pi(\theta)\,d\theta}$$

Notons que dans cette formule $f(x;\theta)$ peut être aussi bien une densité qu'une fonction de probabilité.

On prend alors comme estimation bayesienne $\widehat{\theta}^B$de θ, **la moyenne de la loi a posteriori.** L'estimateur bayésien s'obtient en appliquant à X la fonction associant à une valeur de x quelconque la valeur $\widehat{\theta}^B$ correspondante.

Les avantages de cette approche sont multiples du fait que l'on dispose d'une loi pour θ. Entre autres :

- on peut déterminer aisément un intervalle de valeurs plausibles pour θ.
- on peut estimer θ selon divers critères d'erreur. Le critère des moindres carrés, par exemple, choisit le nombre $\widehat{\theta}^B$ minimisant $E[(\theta - \widehat{\theta}^B)^2]$,où ici θ est aléatoire, ce qui correspond à la moyenne de la loi *a posteriori* .
- on peut estimer toute fonction de θ en calculant directement, pour le critère des moindres carrés, l'espérance de $h(\theta)$ sur la loi a posteriori, soit :

$$E(h(\theta)) = \int_\Theta h(\theta)\,\pi_{\theta/\varkappa=x}(\theta)\,d\theta$$

1.4.2 prédiction Bayesienne

Le modèle statistique associé à ce problème de décision est $\left(\mathcal{X}, \mathcal{B}, P^{\theta}\right)_{\theta \in \theta}$ où $(\mathcal{X}, \mathcal{B})$ est l'espace des observations muni de sa tribu Borélienne ; $\left(P^{\theta}\right)_{\theta \in \Theta}$ une famille de mesures de probabilité sur $(\mathcal{X}, \mathcal{B})$; θ étant l'espace des paramètres.

Soient : $(D, \mathcal{D})$ l'ensemble des décisions muni de sa tribu Borélienne et v une fonction de perte supposée quadratique

$$\begin{array}{rcl} \Theta \times D & \to & \mathbb{R}_{+} \\ v(\theta, d) & = & \int_{R^{+}} (y-d)^{2} f(y/d)\, dy. \end{array}$$

Dans ce cas précis on a : $\mathcal{X} = \{0, 1, ..., n\}$ et $D = \mathbb{R}^{+}$; ceci pour des données de type attribut et $\mathcal{X} = \mathbb{R}^{+n}$ pour des données complètes ou censurées de type II.

On munit l'espace des paramètres Θ d'une mesure de probabilité τ de densité $\pi(\theta)$; la densité à postériori de θ est alors $\pi(\theta/x)$:

$$\pi(\theta/x) = \frac{L(x/\theta)\, \pi(\theta)}{\int_{\Theta} L(x/\theta)\, \pi(\theta)\, d\theta}. \tag{1.20}$$

Ou $L(x/\theta)$ est la fonction de vraisemblance quant on a observé x .On calcule ensuite pour chaque décision $d \in D$ le risque à postériori $R_{\tau}^{x}(d)$

$$\begin{aligned} R_{\tau}^{x}(d) &= \int_{\Theta} v(\theta, d)\, \pi(\theta/x)\, d\theta \qquad (1.21) \\ &= \int_{R^{+}} \int_{\Theta} (y-d)^{2} f(y/\theta)\, \pi(\theta/x)\, d\theta dy \\ &= E^{x}\left((y-d)^{2}\right). \end{aligned}$$

or

$$\pi(y/x) = \int_{\Theta} f(y/\theta)\, \pi(\theta/x)\, d\theta \tag{1.22}$$

est une densité, appelée densité prédictive de Y quand x est observée.
On cherche $\min_{d \in D} R_{\tau}^{x}(d)$; celui-çi est atteint pour

$$\min_{d \in D} R_{\tau}^{x}(d) = V^{x}(Y)$$

On conclut que la règle de décision δ_b de Bayes associe à tout x $\in D$, un prédicteur $\delta_b(x)$ tel que :

$$\delta_b(x) = E^{x}(y) = E^{x}\left(E^{\theta}(Y)\right) = \int_{\Theta} E^{\theta}(Y)\, p(\theta/x)\, d\theta. \tag{1.23}$$

1.4.3 Modélisation de l'information a priori

Le plus souvent on ne dispose pas suffisamment d'information a priori sur le paramètre inconnu θ pour construire la loi a priori. Dans la pratique on a recours à des lois usuelles (loi normales, loi gamma, etc) ou à des lois dites conjuguées (voir ci-dessous), l'information a priori étant alors utilisée pour déterminer les paramètres de la loi a priori, appelés hyperparamètres. En l'absence d'information a priori on introduira la notion de loi a priori non informative qui permet de rester dans un cadre bayésien, alors même que l'on ne dispose pas d'information a priori.

Lois a priori conjuguées

Definition 10 *La loi des observations étant supposée connue, on se donne une famille $\mathcal{F}$ de lois de probabilité sur Θ.On suppose que la loi a priori apartient à $\mathcal{F}$. Si dans ces conditions, la loi a posteriori appartient encore à $\mathcal{F}$, on dit que la loi a priori est conjugée.*

Lois a priori non informatives

Partons d'un exemple pour introduire la notion de loi a priori non informative. On considère le modèle statistique bayésien suivant : les $X_i|\theta$ sont i.i.d. et suivent une loi de Bernoulli de paramètre $\theta \in \left]0,1\right[$. Il n'y a pas une unique loi a priori non informative pour le paramètre θ. On peut en fait proposer différentes lois a priori non informatives.

- En l'absence d'information a priori sur θ, il est naturel de proposer une loi uniforme sur θ car elle donne une probabilité égale aux intervalles de longueur l donnée, à savoir l.

- On peut également proposer la loi a priori (impropre) de Haldane $\pi(\theta) = [\theta(1-\theta)]^{-1} 1_{[0,1]}(\theta)$,en arguant que $E[\theta|x]$ est égal à l'estimateur du maximum de vraisemblance

- une alternative a été proposée par Jeffreys en 1960.

Definition 11 *Soit θ.un paramètre réel. On appelle loi a priori non informative de Jeffreys, la loi 'éventuellement impropre) de densité*

$$\pi_J(\theta) \propto [I(\theta)]^{\frac{1}{2}} \tag{1.24}$$

Definition 12 *où $I(\theta)$ désigne l'information de Fisher apportée par $\boldsymbol{x}$ sur θ.*

1.5 Fonctions de perte

1.5.1 Fonction de perte quadratique

Proposée par Legendre (1805) et Gauss (1810), cette perte est sans aucun doute le critère d'évaluation le plus commun.

$$L(\theta, d) = (\theta - d)^2$$

Dans son article de 1810, Gauss a déjà reconnu le caractère aléatoire de la perte quadratique et le défendait pour des raisons de simplicité. De telles critiques restent valides aujourd'hui. Mais cette perte n'en demeure pas moins utilisée car elle donne en général des solutions bayésiennes acceptables, i.e. celles fournies par une inférence non-décisionnelle fondée sur la densité a posteriori.

Proposition 13 *L'estimateur de Bayes δ^π associé à la distribution a priori π et avec la perte quadratique, est donné par l'espérance a posteriori :*

$$\delta^\pi(x) = E^\pi[\theta/x] = \frac{\int_\theta \theta f(x/\theta)\,\pi(\theta)\,d\theta}{\int_\theta f(x/\theta)\,\pi(\theta)\,d\theta}$$

Corollary 14 *Quand $\Theta \subseteq \mathbb{R}^p$, l'estimateur de Bayes δ^π associé avec π et avec la perte quadratique pondérée $L(\theta, \delta) = (\theta - \delta)^t Q(\theta - \delta)$ est l'espérance a posteriori*

$$\delta^\pi(x) = E^\pi[\theta/x],$$

pour chaque matrice Q,de dimension $p \times p$,symétrique et positive définie.

1.5.2 Fonction de perte erreur absolue

Une solution alternative à la perte quadratique en, une dimension est l'utilisation de la perte erreur absolue :

$$L(\theta, d) = |\theta - d|,$$

considérée par Laplace (1773), ou, plus généralement une fonction multilinéaire :

$$L_{k_1,k_2}(\theta, d) = \begin{cases} k_2(\theta - d), & \text{si } \theta \succ d \\ k_1(d - \theta) & \text{sinon} \end{cases}$$

Ces fonctions croissent moins vite que les fonctions de perte quadratique.Huber (1964) propose un mélange entre les fonctions de pertes erreur absolue et les fonctions de perte quadratique, de manière à garder la pénalisation quadratique autour de 0.

$$L(d,\theta)=\begin{cases}(d-\theta)^2, & \text{si } |d-\theta|\prec k\\ 2k|d-\theta|-k^2, & \text{sinon}\end{cases}$$

Proposition 15 *Un estimateur de Bayes associé à la distribution π et à la perte multi-linéaire, est un $\left(\frac{k_2}{k_1+k_2}\right)$ fractile de $\pi(\theta\backslash x)$.*

Demonstration :Nous avons

$$E(L(\theta,a)/x)$$

$$=k_1\int_{-\infty}^{a}(a-\theta)\,\pi(\theta\backslash x)\,d\theta+k_2\int_{a}^{+\infty}(\theta-a)\,\pi(\theta\backslash x)\,d\theta$$

$$=k_1\int_{-\infty}^{a}P(\theta\prec y\backslash x)\,dy+k_2\int_{a}^{+\infty}P(\theta>y\backslash x)\,dy$$

par une integration par parties.Dérivant par rapport à a,nous obtenons alors

$$k_1P(\theta\prec a\backslash x)-k_2P(\theta>a\backslash x)=0,$$

soit

$$P(\theta\prec a\backslash x)=\frac{k_2}{(k_1+k_2)}$$

En particulier, si $k_1=k_2$ c'est-à-dire en cas d'une erreur de perte absolue, l'estimateur de Bayes est la médiane a posteriori, qui est l'estimateur obtenu par Laplace. Il est à noter que lorsque π a un support discontinu, la proposition précédente donne un exemple d'estimateur de Bayes multiple pour certaines valeurs de x.

1.5.3 Fonction de Perte 0-1

Cette fonction de perte est utilisée surtout dans les méthodes classiques de test d'hypothèse, formalisées par Neyman et Pearsom Plus généralement, c'est un exemple typique de perte non quantitative. En effet, pour cette perte, la pénalité associée avec un estimateur δ est 0 si la réponse est correcte et 1 sinon.

Proposition 16 *.L'estimateur de Bayes associé avec π et avec la perte (2.3.3) est*

$$\delta^{\pi}(x) = \begin{cases} 1, & \text{si } P(\theta \in \Theta_0 \backslash x) \geq P(\theta \notin \Theta_0 \backslash x) \\ 0, & \text{sinon} \end{cases}$$

donc δ^{π} égal à 1 si et seleument si $P(\theta \in \Theta_0 \backslash x) \geq \frac{1}{2}$.

1.5.4 Fonction de Perte intrinsèque

Il arrive que certains cas soient si non-informatifs que non seulement la fonction de perte est inconnue, mais il n'y a même pas de paramétrisation naturelle. Ces cas arrivent lorsque la distribution $f(x\backslash\theta)$ est elle-même d'intérêt, par exemple, dans les cas de prédictions. Dans ces cas non-informatifs, il paraît naturel d'utiliser des fonctions de perte qui comparent directement les distributions $f(.\backslash\theta)$ et $f(.\backslash\delta)$ associées avec le vrai paramètre θ et l'estimé δ. La fonction de perte

$$L(\theta, \delta) = d(f(.\backslash\theta), f(.\backslash\delta))$$

est en fait sans paramètres.Deux mesures de distances usuelles sont :

1-L'entropie :$L_e(\theta - \delta) = E\left[\log\left(\frac{f(x\backslash\theta)}{f(x\backslash\delta)}\right)\right]$qui est aussi appelée la divergence de Kullback-Leibler et qui n'est pas une distance dans le sens mathématique,car elle n'est pas symétrique.

2-La distance de Hellinger :$L_H(\theta, \delta) = \frac{1}{2}E_\theta\left[\left(\sqrt{\frac{f(x\backslash\backslash\delta)}{f(x\backslash\backslash\theta)}} - 1\right)^2\right]$.

1.5.5 Fonction de perte Linex

Une fonction de perte asymétrique très pratique est la fonction de perte Linex (Linear Exponential). Elle a été introduite par Varian (1975). Cette fonction croît presque exponentiellement d'un coté de zéro et est approximativement linéaire de l'autre côté. Sous l'hypothèse que la perte minimale est obtenue pour $\grave{u} = u$, la fonction de perte Linex pour u soit α, soit β s'exprime par

$$L(\Delta) \propto e^{a\Delta} - a\Delta - 1, \quad a \neq 0(+)$$

Où $\Delta = (\grave{u} - u)$ et $\grave{u}$ est un estimateur de u.Le signe et la norme de a représentent respectivement la direction et le degré de symétrie ($a \succ 0$:la surestimation est plus grav
a proche de zéro,la perte Linex est approximativement la fonction de perte quadratique ;(+) devient :

$$E_u(L(\grave{u} - u)) \propto e^{a\grave{u}} E_u(e^{-au}) - a(\grave{u} - E_u(u)) - 1$$

où $E_u(.)$ représente l'espérance a posteriori relative à la densité a posteriori de u. l'estimateur de bayes $\grave{u}_L$ de $\grave{u}$ sous la fonction de perte linex est la valeur de $\grave{u}$ qui minimise (2.3.8)

Pour trouver l'estimateur ;nous dérivons l'équation (2.3.8) par rapport à $\grave{u}$ nous obtenons :

En égalant cette expression à 0, nous obtenons

$$ae^{-a\grave{u}}E_u\left(e^{-au}\right) = a,$$

d'où

$$e^{-a\grave{u}} = E_u\left(e^{-au}\right)$$

En appliquant le logarithme,nous trouvons

$$-a\grave{u} = \log E_u\left(e^{-au}\right)$$

Alors,l'estimateur de Bayes $\grave{u}_L$ de $\grave{u}$ sous la fonction de perte Linex est :

$$\grave{u}_L = -\frac{1}{a}\log(E_u(e^{-au}))E_u(e^{-au}),$$

Etant donné que $E_u(e^{-au})$ existe et est fini.

1.6 Statistique d'ordre

Cette notion est très utile dans une série de problèmes, notamment ceux de minima et de maxima .Comme précédemment nous considérons un échatillon $X_1, X_2, ..., X_n$ dont la loi mère a pour fonction F.

Pour une série de nombres réels $(x_1, x_2, ..., x_n)$ notons $max\{x_1, x_2, ..., x_n\}$ la fonction de $\mathbb{R}^n$ dans $\mathbb{R}$ qui lui associe le nombre maximale de cette série. On peut donc définir une v.a, notée $X_{(n)}$, fonction de $(X_1, X_2, ..., X_n)$ par :

$$X_{(n)} = max\{x_1, x_2, ..., x_n\}$$

La fonction de répartition de cette statistique se déduit aisément de F. En effet l'événement $\left(X_{(n)} \leq x\right)$ est équivalent à $(X_1 \leq x, X_2 \leq x, ..., X_n \leq x)$. Par conséquent :

$$\begin{aligned} F_{X_{(n)}}(x) &= P(X_1 \leq x, X_2 \leq x, ..., X_n \leq x) \\ &= P(X_1 \leq x)P(X_2 \leq x)...P(X_n \leq x) \text{ (indépendance)} \\ &= [F(x)]^n \text{ (même loi)} \end{aligned}$$

De façon similaire on note $X_{(1)} = \min\{X_1, X_2, ..., X_n\}$ la fonction minimum et, en notant que l'événement $\left(X_{(1)} > x\right)$ est équivalent au fait que toutes les X_i sont supérieures à x, on a :

$$\begin{aligned} P\left(X_{(1)} > x\right) &= P(X_1 > x) P(X_2 > x) ... P(X_n > x) \\ &= [1 - F(x)]^n, \end{aligned}$$

d'où :

$$F_{X_{(1)}}(x) = 1 - [1 - F(x)]^n.$$

Definition 17 *Soit h_k la fonction de $\mathbb{R}^n$ dans $\mathbb{R}$ qui à $(x_1, x_2, ..., x_n)$ fait correspondre la k-ième valeur parmi $x_1, x_2, ..., x_n$ lorsqu'on les range dans l'ordre croissant.*
*On appelle alors **statistique d'ordre** k, la v.a. notée $X_{(k)}$,définie par :*

$$X_{(k)} = h_k(X_1, X_2, ..., X_n)$$

Ceci généralise les notions de minimum $(k = 1)$ et de maximum $(k = n)$.

Proposition 18 *La fonction de répartition de $X_{(k)}$ est :*

$$F_{X_{(k)}}(x) = \sum_{j=k}^{n} \binom{n}{j} [F(x)]^j [1 - F(x)]^{n-j}$$

Pour montrer cela il suffit de noter que l'événement $\left\{X_{(k)} \le x\right\}$ est équivalent au fait qu'au moins k v.a. parmi $X_1, X_2, ..., X_n$ soient inférieures à k.Soit X la v.a. symbolisant la loi mère. Considérons l'expérience de Bernoulli avec pour "succès" l'évenement $(X \le x)$ dont la probabilité est $F(x)$. Le nombre de v.a. parmi $X_1, X_2, ..., X_n$ prenant une valeur inférieure ou égale à x est donc une v.a. de loi binomiale $B(n, F(x))$. Pour obtenir la probabilité que ce nombre soit au moins égale à k on est amené à sommer les termes de cette binomiale de k à n.

Remark 19 *Considérons $X_{(i)}$ et $X_{(j)}$ avec $i < j$. La v.a $U = X_{(j)} - X_{(i)}$ ne peut prendre que des valeurs positives et donc $P(U \ge 0) = 1$.De façon conventionnelle on écrira $P\left(X_{(j)} \ge X_{(i)}\right) = 1$ et même , de façon quelque peu rapide, $X_{(j)} \ge X_{(i)}$. Moyennant cette convention, il est possible, comme dans la plupart des ouvrages, de définir les statistiques d'ordre $X_{(1)}, X_{(2)}, ..., X_{(n)}$ par une permutation de $(X_1, X_2, ..., X_n)$ telle que $X_{(1)} \le X_{(2)} \le ... \le X_{(n)}$.*

1.7 Données censurées

Definition 20 *La variable de cencure C est définie par la possible non-observation de l'événement. Si l'on observe C, et non X, et que l'on sait que $X > C$ respectivement $(X < C, C_1 < X < C_2)$, on dit qu'il a censure à droite (résrespectivement censure à gauche, censure par intervalle).*

Si l'événement se produit, X est "réalisée". S'il ne se produit pas (l'individu étant perdu de vue, ou bien exclu vivant), c'est C qui est "réalisée".

X peut être considérée comme la durée séparant un événement initial A d'un événement terminal B, ou comme la durée pendant laquelle un sujet reste dans un état donné (auquel cas A désigne l'entrée dans cet état et B la sortie de cet état - par exemple le chômage). La censure à droite, dont il sera essentiellement question par la suite, est due à la non-observation de B, dont on sait seulement qu'il sera postérieur à la dernière date d'observation du sujet.

Par ailleurs, la censure se distingue de la troncature : on dit qu'il y a troncature à droite (respectivement à gauche) lorsque la variable d'intérêt X_i (durée de vie du i^e individu) n'est pas observable quand elle est supérieure (respectivement inférieure) à un seuil $c > 0$ fixé.

Dans le cas de la censure, on sait que la variable X non observée est supérieure ou inférieure à une valeur C qui, elle, a été observée. La troncature, quand à elle, élimine de l' étude une partie des X_i, ce qui a pour conséquence de faire porter l'analyse uniquement sur la loi de X conditionnellement à l'événement $\{X < c\}$ (respectivement $\{X > c\}$).

Enfin, le mécanisme de censure est habituellement supposé être indépendant de l'événement étudié : on parle de censure non-informative (ignorable). En pratique, cela veut dire que les individus ne doivent pas être censurés parce qu'ils ont un risque de décès pariculièrement élevé (ou faible). En d'autres termes, les individus exclus- vivants ou perdus de vue à une date t doivent être représentatifs des individus encore à risque à cet instant t.

Si la censure est informative, alors l'expression classique de la vraisemblance ne correspond plus à une vraisemblance complète, mais à une vraisemblance partielle qui peut être utilisée pour des inférences, bien qu'il y ait une perte d'efficacité des estimateurs produits (car toute l'information n'est pas utilisée). Ainsi, la censure informative est à l'origine d'un biais lors de l'analyse standard basée sur la vraisemblance (Kalbfleisch et Prentice, 1980 ; Schluchter, 1992).

Definition 21 *On dispose de données complètes quand le test de survie est mené jusqu'à la n^{ieme} panne. Le vecteur des observations est alors $x = (x_1, x_2, ...x_n)$;où x_i est la durée de vie de i^{ieme} item.*

La vraisemblance associée à un tel plan d'expérience est alors :

$$L(x|\theta) = \prod_{i=1}^{n} f(x_i|\theta) \tag{1.25}$$

ce plan est noté $[n, B, n]$.

Definition 22 *La censure est dite* ***non aléatoire de type I*** *si, étant donné un n-échantillon $X_1.....X_n$, d'une variable aléatoire X, dépendant d'un paramètre θ, il existe une v.a. n-dimensionnelle $(C_1, ..., C_n\ )$ de $\mathbb{R}^{+n}$ telle que les observations consistent en (T_i, δ_i), où*

$$\begin{cases} T_i = X_i \wedge C_i \\ \delta_i = \mathbf{1}_{\{X_i \leq C_i\}} \end{cases}$$

Definition 23 *La censure est dite de* ***type II*** *si, étant donné un nombre positif fixé r et un n-échantillon $X_1.....X_n$, les observations consistent en (T_i, δ_i), où*

$$\begin{cases} T_i = X_{(i)} \wedge X_{(r)} \\ \delta_i = \mathbf{1}_{\left\{X_{(i)} \leq X_{(r)}\right\}} \end{cases}$$

où $X_{(1)} < X_{(2)} < ..... < X_{(n)}$ sont les statistiques d'ordre de $X_1.....X_n$.
La vraisemblance est alors :

$$L\left(x/\theta\right) = \frac{n!}{(n-r)!} \prod_{i=1}^{r} f\left(x_i\right) \left[1 - F\left(x_r\right)\right]^{n-r}$$

Definition 24 *On fixe la durée du test à un temps T ; les données issues de ce test sont résumées par le nombre x de pannes ayant eu lieu dans $[0, T]$; x=1,...,n. On dit alors qu'on dispose de données de type attribut.*

La vraisemblance est alors :

$$L\left(x/\theta\right) = \binom{n}{x} \left[F\left(T/\theta\right)\right]^{x} \left[1 - F\left(T/\theta\right)\right]^{n-x}. \tag{1.26}$$

Il est supposé que tout item tombé en panne dans $[0, T]$ est non renouvellé.

Chapitre 2

Modèle Exponentielle

2.1 Introduction

Le problème de la prédiction de statistiques d'ordre et de fonctions de celles-çi et ses applications en fiabilité a fait l'objet de plusieurs articles. Plusieurs auteurs se sont intéressés à ce problème en utilisant une approche classique ; Dunsmore (1974) [21] ; Bancroft G.A. (1976) [2] ; Prezzi (1996) [10] ; Pan JN. (1997) [45]. L'approche Bayesienne a été également utilisée avec des lois à priori conjuguées naturelles ou non informative ; des données complètes ou censurées et différentes lois. Bratcher (1971) [3] a utilisé des données complètes ; Dunsmore (1974) [20] a étudié d'une manière générale les tests de survie ; Evans et Nigm (1980) [23, 24] ont étudié la loi exponentielle tronquée et la loi exponentielle bivariée avec des données censurées de type II ; Arnold BC. (1989) [1] a considéré la loi de Pareto ; Erkanly (1998) [22] et Van-batenberg (1989) [52] ont appliqué les résultats de la prédiction Bayesienne à la fiabilité.

On dispose d'une experience E dans laquelle n items sont soumis à un test de survie.La durée de vie de chaque item est $X_{i\,;}\, i = 1, ..., n$. Les X_i suivent une loi exponentielle à un paramètre :

$$f\left(x/\theta\right) = \theta \exp\left(-\theta x\right); \theta > 0; x \geqslant 0 \qquad (*) \tag{2.1}$$

La fonction de répartition est

$$F(x/\theta) = 1 - \exp(-\theta x); \theta > 0; x \geqslant 0 \tag{2.2}$$

On s'intéresse d'une part à la prédiction dans un futur échantillon $(Y_{1,}..., Y_N)$ indépendant du premier de

$$Y = Y_{(k)}; 1 \leq k \leqslant N \tag{2.3}$$

$$Y = \sum_{i=1}^{k} Y_{(i)} + (N-k)\, Y_{(k)} \tag{2.4}$$

et d'autre part à la prédiction dans un même échantillon de

$$Y = X_{(k)};\ r < k \leqslant n \tag{2.5}$$

$$Y = \text{nombre de pannes dans } \left[X_{(r)}; X_{(r)} + T\right]$$

où $\left(Y_{(1)}, Y_{(2)}, ..., Y_{(k)}\right)$ est la statistique d'ordre de $(Y_1, Y_2, ..., Y_N)$; X_r la r^{ieme} statistique d'ordre ; T>0 est un paramètre temps fixé d'avance.

On se propose le même problème pour des données de type attribut.On applique les résultats à la durée de vie de quelques systèmes.

2.2 Description du problème et préliminaires

On observe n v.a. $X_1, X_{2,...}X_n$ indépendantes et identiquement distribuées de loi P_0^θ et à valeurs dans $(\mathcal{X}_{0,}\mathcal{B}_0)$. On suppose alors que les observations sont de type attribut pour prévoir (2.1) , (2.2) , (2.3) et (2.4) .

2.2.1 Données de type attribut

On fixe la durée du test à un temps T ; les données issues de ce test sont résumées par le nombre x de pannes ayant eu lieu dans $[0, T]$; x=1,...,n.

La vraisemblance est alors :

$$L(x/\theta) = \binom{n}{x} \left[F(T/\theta)\left[1 - F(T/\theta)\right]\right]^{n-x}.$$

Il est supposé que tout item tombé en panne dans $[0, T]$ est non renouvellé.

2.2.2 Distribution de la k^{ieme}statistiqued$'$ordre

$(Y_1, Y_2, ..., Y_N)$; X_r étant un N-échantillon de même loi que X.

Y=Y$_{(k)}$; 1$\leq$ k $\leq$ N; représente la durée de vie du k^{ieme} item tombé en panne.

$\left(Y_{(1)}, Y_{(2)}, ...Y_{(N)}\right)$ représente la statistique ordonnée.

$$Y = \sum_{i=1}^{k} Y_{(i)} + (N-k)\, Y_{(k)}$$

représente la durée de vie totale des N items jusqu'à la k^{ieme}panne.La densité de Y=Y$_{(k)}$ est :

$$f_k\left(y/\theta\right)=k\binom{N}{k}\left[F\left(y/\theta\right)\right]^{k-1}f\left(y/\theta\right)\left[1-F\left(y/\theta\right)\right]^{N-k};\qquad y\geqslant 0$$

Sa fonction de répartition est :

$$.F_k\left(y/\theta\right)=\sum_{i=k}^{N}\binom{N}{i}\left[F\left(y/\theta\right)\right]^{i}\left[1-F\left(y/\theta\right)\right]^{N-i}$$

2.2.3 Résolution de ce problème de décision par la méthode de Bayes

Le modèle statistique associé à ce problème de décision est $\left(\mathcal{X},\mathcal{B},P^{\theta}\right)_{\theta\in\theta}$ où $\left(\mathcal{X},\mathcal{B}\right)$ est l'espace des observations muni de sa tribu Borélienne ; $\left(P^{\theta}\right)_{\theta\in\Theta}$ une famille de mesures de probabilité sur $\left(\mathcal{X},\mathcal{B}\right)$; θ étant l'espace des paramètres.

Soient : $\left(D,\mathcal{D}\right)$ l'ensemble des décisions muni de sa tribu Borélienne et v une fonction de perte supposée quadratique

$$\begin{array}{rcl}\Theta\times D & \rightarrow & \mathbb{R}_{+}\\ v\left(\theta,d\right) & = & \int_{R^{+}}\left(y-d\right)^{2}f\left(y:d\right)dy.\end{array}$$

Dans ce cas précis on a : $\mathcal{X}=\{0,1,...,n\}$ et $D=\mathbb{R}^{+}$; ceci pour des données de type attribut.

On munit l'espace des paramètres Θ d'une mesure de probabilité τ de densité $p\left(\theta\right)$; la densité à postériori de θ est alors $p\left(\theta/x\right)$:

$$p\left(\theta/x\right)=\frac{L\left(x/\theta\right)p\left(\theta\right)}{\int_{\Theta}L\left(x/\theta\right)p\left(\theta\right)d\theta}.$$

Ou $L\left(x/\theta\right)$ est la fonction de vraisemblance quant on a observé x .On calcule ensuite pour chaque décision $d\in D$ le risque à postériori $R_{\tau}^{x}\left(d\right)$

$$\begin{array}{rcl}R_{\tau}^{x}\left(d\right) & = & \int_{\Theta}v\left(\theta,d\right)p\left(\theta/x\right)d\theta\\ & = & \int_{\mathcal{X}_{0}}\left(y-d\right)^{2}p\left(y/x\right)dy\\ & = & E^{x}\left(\left(y-d\right)^{2}\right).\end{array}$$

or

$$p\left(y/x\right)=\int_{\Theta}f\left(y/\theta\right)p\left(\theta/x\right)d\theta \tag{2.6}$$

est une densité, appelée densité prédictive de Y quand x est observée.
On cherche $\min\limits_{d \in D} R_{\tau}^{x}(d)$; celui-çi est atteint pour

$$\min_{d \in D} R_{\tau}^{x}(d) = V^{x}(Y)$$

On conclut que la règle de décision δ_b de Bayes associe à tout x $\in D$, un prédicteur $\delta_b(x)$ tel que :

$$\delta_b(x) = E^{x}(y) = E^{x}\left(E^{\theta}(Y)\right) = \int_{\Theta} E^{\theta}(Y)\, p(\theta/x)\, d\theta. \tag{2.7}$$

2.3 Prédiction dans le cas ou les données sont de type attribut

Soit x le nombre de défaillances dans $[0, T]$; $x = 0, 1, ...n$; T étant fixé d'avance. On suppose sur θ une densité a priori de type conjuguée naturelle.

2.3.1 Vraisemblance et densité à postériori

La loi du nombre de pannes dans $[0, T]$ est une loi binomiale $\mathcal{B}(n, 1 - \exp(-\theta T))$; d'ou la vraissemblance

$$\begin{aligned} L(x/\theta) &= \tbinom{n}{x} [1 - \exp(-\theta T)]^{x} \exp(-\theta T(n-x)) \\ &= \tbinom{n}{x} \sum_{i=0}^{x} (-1)^{i} \tbinom{x}{i} \exp(-\theta T(n-x+i)). \end{aligned}$$

La loi à priori sur θ étant une loi gamma $\Gamma(g, h)$; la densité à postériori est alors donnée par

$$p(\theta/x) = \frac{1}{A\Gamma(g)} \sum_{i=0}^{x} (-1)^{i} \tbinom{x}{i} \theta^{g-1} \exp(-\theta(H + iT)) ;\ \theta > 0 \tag{2.8}$$

$$\text{où } \Gamma(g) = \int_0^{\infty} x^{g-1} e^{-x} dx.$$

$$H = h + T(n-x).$$

$$A = \sum_{i=0}^{x} (-1)^{i} \tbinom{x}{i} (H + iT)^{-g}.$$

2.3.2 Prédiction de $Y = Y_{(k)}; 1 \leq k \leq N$

La densité de la $k^{\text{ième}}$ statistique d' ordre est donnée par

$$f\left(y/\theta\right) = k\binom{N}{k}\sum_{j=0}^{k-1}\left(-1\right)^{j}\binom{k-1}{j}\theta\exp\left(-\theta y\left(N-k+1+j\right)\right);\ y \geq 0.$$

L'intégration par rapport à θ du produit $f\left(y/\theta\right)p\left(\theta/x\right)$ conduit à la densité prédictive

$$\begin{aligned} p\left(y/x\right) &= \frac{k\binom{N}{k}}{A\Gamma\left(g\right)}\sum_{i=0}^{x}\sum_{j=0}^{k-1}\left(-1\right)^{i+j}\tbinom{x}{i}\binom{k-1}{j} \qquad (2.9)\\ &\quad\times\int_0^\infty \theta^g\exp\left(-\theta\left(y\left(N-k+1+j\right)+H+iT\right)\right)d\theta\\ &= \frac{gk\binom{N}{k}}{A}\sum_{i=0}^{x}\sum_{j=0}^{k-1}\left(-1\right)^{i+j}\tbinom{x}{i}\binom{k-1}{j}\\ &\quad\times\left[y\left(N-k+1+j\right)+H+iT\right]^{-(g+1)}. \end{aligned}$$

La fonction de répartition prédictive de Y est

$$F\left(z/x\right) = \int_0^z p\left(y/x\right)dy$$

$$= \frac{gk\binom{N}{k}}{A}\sum_{i=0}^{x}\sum_{j=0}^{k-1}\left(-1\right)^{i+j}\tbinom{x}{i}\binom{k-1}{j}\times\int_0^z\left[y\left(N-k+1+j\right)+H+iT\right]^{-(g+1)}dy$$

$$\begin{aligned} F\left(z/x\right) &= \frac{k\binom{N}{k}}{A}\sum_{i=0}^{x}\sum_{j=0}^{k-1}\left(-1\right)^{i+j}\binom{x}{j}\binom{k-1}{j}\left(N-k+1+j\right)^{-1} \qquad (2.10)\\ &\quad\times\left[\left(H+iT\right)^{-g}-\left(z\left(N-k+1+j\right)+H+iT\right)^{-g}\right]. \end{aligned}$$

En particulier pour $k = 1$

$$F\left(z/x\right) = 1-\frac{1}{A}\sum_{i=0}^{x}\left(-1\right)^{i}\tbinom{x}{i}\left[Nz+H+iT\right]^{-g}.$$

$p\left(y/x\right)$ étant une fonction stictement décroissante de y, si on cherche un intervalle de prédiction de niveau $(1-\gamma)$ et de la forme (0,a) ; alors a est solution de

$$\frac{1}{A}\sum_{i=0}^{x}\left(-1\right)^{i}\tbinom{x}{i}\left[Na+H+iT\right]^{-g} = \gamma.$$

La solution de l'équation çi-dessus peut être approchée par une méthode itérative.

Prédicteur de $\mathbf{Y}_{(k)}$
Sous l'hypothèse $g > 1$; un prédicteur de Y serait

$$\begin{aligned} E\left(Y/x\right) &= \int_0^\infty yp\left(y/x\right)dy \qquad (2.11) \\ &= \frac{k\binom{N}{k}}{A\left(g-1\right)} \sum_{i=0}^{x}\sum_{j=0}^{k-1} \left(-1\right)^{i+j} \binom{x}{j}\binom{k-1}{j} \frac{\left(N-k+1+j\right)^{-2}}{\left(H+iT\right)^{g-1}} \\ &= \frac{gk\binom{N}{k}}{A\left(g-1\right)} \left[\sum_{i=0}^{x} \left(-1\right)^i \binom{x}{i} \left(H+iT\right)^{1-g}\right] \\ &\quad \times \left[\sum_{j=0}^{k-1} \left(-1\right)^j \binom{k-1}{j} \left(N-k+1+j\right)^{-2}\right]. \end{aligned}$$

En particulier pour $k = 1$:

$$E\left(Y_1/x\right) = \frac{1}{A\left(g-1\right)N}\left[\sum_{i=0}^{x}\left(-1\right)^i\binom{x}{i}\left(H+iT\right)^{1-g}\right].$$

2.3.3 Prédiction de $\mathbf{Y}= \sum_{i=0}^{k} Y_{(i)} + \left(N-k\right)Y_{(k)}$

Y suit une loi $\Gamma\left(k,\theta\right)$ dont la densité est

$$p\left(y/\theta\right) = \frac{\theta^k}{\Gamma\left(k\right)} y^{k-1}\exp\left(-\theta y\right); y \geq 0, \theta > 0.$$

Soit p(y/x), la densité prédictive de Y :

$$\begin{aligned} p\left(y/x\right) &= \int_0^\infty p\left(y/\theta\right)p\left(\theta/x\right)d\theta \qquad (2.12) \\ &= \frac{1}{A\beta\left(k,g\right)}\sum_{i=0}^{x}\left(-1\right)^i\binom{x}{i}\frac{y^{k-1}}{\left[y+H+iT\right]^{k+g}}. \end{aligned}$$

où

$$\beta\left(k,g\right) = \int_0^1 t^{k-1}\left(1-t\right)^{g-1}dt.$$

On remarque que la loi conditionnelle de Y sachant x est une combinaison linéaire de loi β_{II}. Ou la loi β_{II} est définie par : $f(y) = \frac{1}{\beta(k,g)} x \frac{y^{k-1}}{(1+y)^{k+g}}$.

La fonction de répartition prédictive $F(z/x)$ de Y est

$$F(z/x) = \frac{1}{A\beta(k,g)} \sum_{i=0}^{x} (-1)^i \binom{x}{i} \int_0^z \frac{y^{k-1}}{[y+H+iT]^{k+g}} dy. \tag{2.13}$$

Un intervalle de prédiction de Y de niveau $(1-\gamma)$ et de la forme (0, z) est tel que z soit solution de

$$\frac{1}{A} \sum_{g\, i=0}^{x} (-1)^i \binom{x}{i} (H+iT)^{-g} I_{(a_i)}(k,g) = 1-\gamma$$

où $I_{(a_i)}(k,g)$ est la valeur d'une loi beta de type II de paramètres k et g au point $a_i = \frac{z}{z+H+iT}$.

Prédicteur de Y

Sous l'hypothèse $g > 1$; un prédicteur de Y serait

$$\begin{aligned} E(Y/x) &= \frac{1}{A\beta(k,g)} \sum_{i=0}^{x} (-1)^i \binom{x}{i} \int_0^\infty \frac{y^k}{[y+H+iT]^{k+g}} dy \qquad (2.14) \\ &= \frac{\beta(k+1,g-1)}{A\beta(k,g)} \sum_{i=0}^{x} (-1)^i \binom{x}{i} (H+iT)^{1-g} \\ &= \frac{k}{A(g-1)} \sum_{i=0}^{x} (-1)^i \binom{x}{i} (H+iT)^{1-g}. \end{aligned}$$

2.4 Prédiction dans un même échantillon

Ayant observé r pannes dans $[0,T]$ et sous l'hypothèse $r < n$, on s'intéresse à la prédiction de $Y = X_{(r+l)} - T$; $l = 1, 2, ...n - r$ et à celle du nombre de pannes Y dans $[T, T+t]$; où t>0 est un temps fixé à l'avance.

2.4.1 Prédiction de $Y = X_{(r+l)} - T$

La densité de $X_{(r+l)}$ sachant $X_{(r)} \leq T < X_{(r+1)}$ est pour $x > T$:

$$\begin{aligned} f^{(r)}(x/\theta) &= \frac{1}{\beta(l, n-r-l+1)} \left[\frac{F(x/\theta) - F(T/\theta)}{1-F(T/\theta)}\right]^{l-1} \\ &\times \left[\frac{1-F(x/\theta)}{1-F(T/\theta)}\right]^{n-r-l} \frac{[f(x/\theta)]}{[1-F(T/\theta)]} \end{aligned}$$

Ce qui donne pour $Y = X_{(r+l)} - T$:

$$f^{(r)}(y/\theta) = l\binom{n-r}{l}\sum_{j=0}^{l-1}(-1)^j\binom{l-1}{j}\theta\exp\left(-\theta y\left(n-r-l+j+1\right)\right); y \geq 0.$$

Densité prédictive de Y : la densité à priori sur θ est supposée celle d'une loi $\Gamma(g,h)$; la densité prédictive $p(y/r)$ est alors :

$$\begin{aligned} p(y/r) &= \int_0^\infty f^{(r)}(y/\theta)\,p(\theta/r)\,d\theta \\ &= \frac{l\binom{n-r}{l}}{A\Gamma(g)}\sum_{i=0}^{r}\sum_{j=0}^{l-1}(-1)^{i+j}\binom{r}{i}\binom{l-1}{j} \\ &\quad\times\int_0^\infty \theta^g\exp\left(-\theta\left(y\left(n-r-l+j+1\right)+H+iT\right)\right)d\theta \\ &= \frac{gl\binom{n-r}{l}}{A}\sum_{i=0}^{r}\sum_{j=0}^{l-1}(-1)^{i+j}\binom{r}{i}\binom{l-1}{j} \\ &\quad\times\left[y\left(n-r-l+j+1\right)+H+iT\right]^{-(g+1)}. \end{aligned} \tag{2.15}$$

En particulier pour $l = 1$:

$$p(y/r) = \frac{g(n-r)}{A}\sum_{i=0}^{r}(-1)^i\binom{r}{i}\left[y(n-r)+H+iT\right]^{-(g+1)}.$$

La fonction de répartition prédictive $F(z/r)$ de Y est alors

$$\begin{aligned} F(z/r) &= \frac{gl\binom{n-r}{l}}{A}\sum_{i=0}^{r}\sum_{j=0}^{l-1}(-1)^{i+j}\binom{r}{i}\binom{l-1}{j} \\ &\quad\times\int_0^z\left(y\left(n-r-l+j+1\right)+H+iT\right)^{-(g+1)}dy \\ &= \frac{l\binom{n-r}{l}}{A}\sum_{i=0}^{r}\sum_{j=0}^{l-1}(-1)^{i+j}\binom{r}{i}\binom{l-1}{j}\left(n-r-l+j+1\right)^{-1} \\ &\quad x\left[\left(H+iT\right)^{-g}-\left(z\left(n-r-l+j+1\right)+H+iT\right)^{-g}\right]. \end{aligned} \tag{2.16}$$

En particulier, pour $l = 1$;un intervalle de prédiction de niveau $(1-\gamma)$ pour $Y = X_{(r+1)} - T$ est tel que z soit solution de :

$$\begin{aligned} F(z/r) &= (1-\gamma) \\ \frac{1}{A}\sum_{i=0}^{r}(-1)^i\binom{r}{i}\left[(n-r)z+H+iT\right]^{-g} &= \gamma. \end{aligned}$$

Prédicteur de $Y = X_{(r+l)} - T$:
Par rapport à une fonction de perte quadratique et sous l'hypothèse $g > 1$

$$\begin{aligned}
E(Y/r) &= \int_0^\infty yp(y/r)\,dy \qquad (2.17)\\
&= \frac{gl\binom{n-r}{l}}{A}\sum_{i=0}^{r}\sum_{j=0}^{l-1}(-1)^{i+j}\binom{r}{i}\binom{l-1}{j}\\
&\quad\times\int_0^\infty \frac{y}{(y(n-r-l+j+1)+H+iT)^{g+1}}dy\\
&= \frac{l\binom{n-r}{l}}{A(g-1)}\sum_{i=0}^{r}(-1)^i\binom{r}{i}[H+iT]^{-(g-1)}\\
&\quad\times\sum_{j=0}^{l-1}(-1)^j\binom{l-1}{j}(n-r-l+j+1)^{-2}.
\end{aligned}$$

En particulier pour $l = 1$

$$E(Y_1/r) = \frac{1}{A(g-1)(n-r)}\sum_{i=0}^{r}(-1)^i\binom{r}{i}[H+iT]^{-g}.$$

2.4.2 Prédiction du nombre de pannes dans $[T, T+t]$ quand on a observé r pannes dans $[0, T]$

Soit Y la variable aléatoire donnant le nombre de pannes dans $[T, T+t]$. De $[Y > l] \Leftrightarrow \left[X_{(r+l+1)} - T \le t\right]$, on a

$$P[(Y > l)\ /r] = P\left[\left(X_{(r+l+1)} - T \le t\right)/r\right].$$

On connait la loi de $X_{(r+l)}$ sachant r ; on peut en déduire la loi de $X_{(r+l+1)}$ sachant r :

$$\begin{aligned}
P[Y > l/r] &= \frac{(l+1)\binom{n-r}{l+1}}{A}\sum_{i=0}^{r}\sum_{j=0}^{l}(-1)^{i+j}(n-r-l+j)^{-1} \qquad (2.18)\\
&\quad\left[(H+iT)^{-g} - (t(n-r-l+j)+H+iT)^{-g}\right].
\end{aligned}$$

En particulier, la probabilité d'avoir au moins une panne dans $[T, T+t]$ est :

$$P[(Y > 0)\ /r] = 1 - \frac{1}{A}\sum_{i=0}^{r}(-1)^i\binom{r}{i}[t(n-r)+H+iT]^{-g}.$$

Un prédicteur de Y serait le plus proche entier de $E(Y/r)$:

$$
\begin{aligned}
E(Y/r) &= \sum_{l=0}^{n-r} P[Y>l] \qquad (2.19)\\
&= \frac{1}{A}\sum_{l=0}^{n-r}(l+1)\binom{n-r}{l+1}\sum_{i=0}^{r}\sum_{j=0}^{l}(-1)^{i+j}\binom{r}{i}\binom{l}{j}(n-r-l+j)^{-1}\\
&\quad\times\left[(H+iT)^{-g}-(t(n-r-l+j)+H+iT)^{-g}\right].
\end{aligned}
$$

2.5 Application des résultats obtenus à la prédiction de la durée de vie de quelques systèmes

Soit S_k un système à k composants identiques ayant chacun une durée de vie dont la densité $f(x/\theta)$ est donnée par $(*)$. A partir d'observations censurées de type II ou d'observations de type attribut d'un n-échantillon de $(*)$, les résultats obtenus en (3.2) et (3.3) permettent la prédiction de la durée de vie du système $S_{k\,.}$

2.5.1 *"R out of k system"*

S_k fonctionne si et seulement si r des k composants fonctionnent simultanément.

"k out of k " correspond à un système de composant en série.

"1 out of k "correspond à un système de composant en parallèle.

On peut considérer que les durées de vie $Y_{1,}Y_{2,}...,Y_k$ de chacun des composants du système constituent un k-échantillon de $(*)$ indépendant de l'échantillon $(X_{1,}X_2,...X_n)$ des observations. Ainsi la prédiction de Y est équivalente à la prédiction de $\text{Y}_{(r)}; 1 \leq r \leq k$. On peut donc appliquer les résultats de (2.11) dans le cas de données de type attribut.

2.5.2 "Stand-by redundant system"

S_k est un système à k composants identiques ; un composant d'origine et $(k-1)$ composants de rechange, chacun commençant son fonctionnement dès que le composant précédent tombe en panne. Ainsi le système S_k est défaillant quand il n'y a plus d'unités de rechange, sa durée de vie Y est donnée par :

$$Y = Y_1 + Y_2 + ... + Y_k$$

On peut donc appliquer les résultats de (2.14) avec $n = k$ et $r = k$

2.5.3 Système en série avec "stand-by composants"

Un tel système est constitué de k composants opérant en série avecc $(r-1)$ unités de rechange. On suppose que chaque unité tombée en panne est remplacée instantanément par une unité de rechange de sorte que la durée de vie Y de S_k soit égale à $Y_{(l)}; 1 \leq l \leq r$. On peut donc appliquer dans ce cas les résultats de (2.14) et (2.15).

2.6 Simulation

On effectue des simulations pour obtenir des prédicteurs et des intervalles de prédiction de $Y_{(1)}$ dans le cas d'une distribution $\exp(\theta)$ pour des données de type attribut. Pour différentes valeurs de θ, on calcule la fonction de risque $R(\theta, d(x))$ en utilisant la loi des grands nombres où :

$$R(\theta, d(x)) = E_\theta\left[(Y - d(x))^2\right].$$

Un prédicteur de $Y = Y(1)$ quand x est observé est obtenu en (2.11).

Un prédicteur de $Y = \sum_{i=1}^{k} Y_{(i)} + (N-k) Y_{(k)}$ quand x est observé est obtenu en (2.14)

Une borne supérieure de prédiction de $Y_{(1)}$ de la forme $(0, z)$ et de niveau $(1-\gamma)$ est telle que z soit solution de :

$$\frac{\sum_{i=0}^{x} (-1)^i \binom{x}{i} [Nz + h + T(n-x) + iT]^{-g}}{\sum_{i=0}^{x} (-1)^i \binom{x}{i} [h + T(n-x) + iT]^{-g}} = \gamma$$

Cette équation est résolue par la méthode de Newton. Les valeurs choisies pour effectuer les simulations sont :

n : nombre d'items observés dans $[0, T]$; n=10.

T : durée de l'observation ; T=10.

g, h : sont les paramètres de la loi à priori sur θ qui est de type $\Gamma(g, h)$; $g = 2$ et $h = 0.01$.

N : taille d'un échantillon futur tiré d'une même loi $\exp(\theta)$, N=20.

x : le nombre de défaillances (noté ix) observés dans $[0, T]$; x=0,1,...,n Les valeurs de x sont simulées.

Les résultats pour chaque valeur de θ sont groupés par tableau :

1ère colonne : les valeurs de ix (simulées)

2ème colonne : les valeurs de la première statistique d'ordre $Y_{(1)}$ d'un échantillon futur indépendant du premier et de taille N (simulées)

3ème colonne : les valeurs pour chaque ix observé d'un prédicteur d(x) de $Y_{(1)}$ (calculées en utilisant un programme fortran).

4ème colonne : on calcule les écarts quadratiques $\left(Y_{(1)} - d\left(x\right)\right)^2$

5ème colonne : pour chaque ix observé, on calcule une borne supérieure de prédiction z de niveau 0.95 et avec une précision eps=0.001.(z est calculée en utilisant la méthode de Newton).

On calcule au bas de chaque tableau :

-la moyenne empirique ME de $Y_{(1)}$

-la moyenne théorique MT de $Y_{(1)}$ (MT=$\frac{1}{N\theta}$)

-la moyenne des prédicteurs MP.

-la valeur de la fonction de risque au point θ :

$$R\left(\theta, d\right) = \frac{1}{N}\sum_{i=1}^{N}\left[Y_{(1)} - d\left(x_i\right)\right]^2$$

Les résultats de la simulation pour $\theta = 0.0142$ sont données ci-dessous :

ix	y(1)	d(x)	$(y(1)-d(x))^2$	z
0	5.70	4.89	0.50	17.36
5	2.07	0.58	2.19	1.90
1	6.19	2.36	14.66	8.12
2	0.58	1.48	0.81	4.07
2	1.16	1.48	0.10	4.97
1	1.18	2.36	1.39	8.12
2	1.21	1.48	0.07	4.97
0	0.40	4.99	21.06	17.36
2	1.29	1.48	0.03	4.97
1	0.07	2.36	5.24	8.12
0	6.60	4.99	2.59	17.36
1	3.33	2.36	0.94	8.12
4	0.74	0.77	0.00	2.51
0	5.06	4.09	0.00	17.36
4	0.74	0.77	0.00	2.51
0	7.16	4.09	4.70	17.36
1	1.29	2.36	1.14	8.12
0	9.53	4.99	20.61	17.36
1	0.10	2.36	5.10	8.12
4	0.85	0.77	0.00	2.51

2.7 Prédiction dans le cas ou les données sont groupées

Nous nous proposons, dans ce travail, de considérer les problèmes de prédictions de statistiques d'ordre et de fonctions de celles-ci dans un échantillon futur de taille N, dans le cas où les observations sur le premier échantillon sont issues d'un plan d'expérience avec des données groupées puis, nous considérons le cas d'un plan d'expérience avec renouvellement . Il est supposé que les durées de vie suivent une loi exponentielle à un paramètre dont la densité est donnée par :

$$f\left(x/\theta\right) = \theta \exp\left(-\theta x\right); \theta > 0; x > 0 \tag{2.20}$$

Soit $\left(X_1, X_2, ..., X_N\right)$ un n-échantillon d'une loi $\exp(\theta)$.Nous disposons d'une experience E dans laquelle n items sont soumis à un test de survie.Dans les deux plans d'experiences cités plus haut ;on s'intéresse à la prédiction dans un échantillon futur $\left(Y_1, Y_2, ..., Y_N\right)$ indépendant du premier des deux variables d'intérét suivantes :

$$\begin{aligned} Y &= Y_{(r)}; 1 < r < N. & (2.21) \\ Z &= \sum_{i=1}^{k} Y_{(i)} + \left(N-k\right) Y_{(k)} & (2.22) \end{aligned}$$

Où$\left(Y_{(1)}, Y_{(2)}, ..., Y_N\right)$ représente la statistique d'ordre de , Y représente la ième statistique d'ordre et Z, la durée de vie globale des N items jusqu'à la kième panne.

Dans le paragraphe 2 , outre la densité prédictive, un prédicteur de(2.21) a été trouvé avec des données groupées. Dans le paragraphe 3 ; la densité prédictive et un prédicteur de(2.21)et de(2.22)ont été trouvés dans le cas d'un plan d'expérience avec renouvellement. 2.

2.7.1 Cas ou les données sont groupées

Considérons un plan d'expérience de durée T fixée, comprenant à l'origine n items soumis à un test de survie. Les durées de vie sont supposées exponentielles de densité donnée en (2.20). L'intervalle de temps [0, T] est subdivisé en k intervalles de longueur .T/k

Soit $\left(t_1, t_2, ..., t_k\right)$ tel que $t_j = jT/k$
$j = 1, ..., k$ des temps d'inspections prédéterminés, avec$t_k = T$, le dernier

instant d'inspection

On suppose $t_0 = 0$.et $t_{k+1 \rightarrow +\infty}$

L'information recueillie à l'issue du test est résumée par les nombres xi de défaillances ayant eu lieu dans chacun des k intervalles$[t_{j-1}, t_j]; j = 1, \ldots, k$.

Soient :

p_1= probabilité de panne dans $\left[0, \frac{T}{K}\right] = F_\theta\left(t_1\right)$

p_j = probabilité de panne dans$\left[t_{j-1}, t_j\right] = F_\theta\left(t_1\right) - F_\theta\left(t_{1-1}\right)$pour $j = 1, \ldots, k$

En vertu de la propriété de perte de mémoire de la loi exponentielle, les probabilités $p_j$vérifient la relation de récurrence suivante :

$$p_j = p_1 \exp\left\{-\theta\left(j-1\right) T/k\right\}$$

et

$$\sum_{j=1}^{k} p_j = F_\theta\left(T\right) \tag{2.23}$$

En effet

$p_j = F_\theta\left(t_j\right) - F_\theta\left(t_{_{j-1}}\right)$

$= \exp\left(-\theta t_{j-1}\right) - \exp\left(-\theta t_j\right)$

$= \exp-\left(\theta\left(j-1\right) T/k\right) - \exp\left(-\theta\left(j\right) T/k\right)$

$= \exp-\left(\theta\left(j-1\right) T/k\right)\left[1 - \exp-\left(\theta T/k\right)\right]$

$= p_1 \exp-\left(\theta\left(j-1\right) T/k\right)$

Vraisemblance et densité à posteriori

Soit

$\underline{x} = \left(x_1, ..., x_k\right)$ le vecteur des observations.

On pose,

$$S = \sum_{i=1}^{k} x_i, S' = nk - \sum_{i=2}^{k}\left(k-i+1\right) x_i; C = \frac{n!}{\prod_{i=1}^{k} x_i!\left(n-s\right)!}$$

La vraisemblance suit alors une loi multinomiale

$$M_{k+1}\left(n, p_1, ..., p_k\right)$$

$$L\left(\underline{x}/\theta\right) = C\prod_{j=1}^{k} p_j^{xj}\left\{1 - \sum_{i=1}^{k} p_1\right\}^{n-s}$$

La relation de récurrence(2.23) nous permet d'écrire la vraisemblance uniquement en fonction de p_1 :

$$L\left(\underline{x}/\theta\right) = Cp_i^s \exp\left\{-\theta\frac{T}{K}x_2 + 2x_3 + ... + (K-1)\,x_k\right\}\left(1 - \sum_{i=1}^{k} p_i\right)^{n-s}$$

$$L\left(\underline{x}/\theta\right) = C\sum_{i=0}^{S}(-1)^i\begin{pmatrix} S \\ i \end{pmatrix}\exp\left\{-\theta\tfrac{T}{K}\left(S' + i\right)\right\}$$

On suppose surθ une distribution à priori de type Gamma,$G\left(g, h\right)$,en effet celle-ci est une conjuguée naturelle, la densité a posteriori est alors :

$$p\left(\theta/\underline{x}\right) = \frac{1}{K\Gamma\left(g\right)}\sum_{i=0}^{S}(-1)^i\begin{pmatrix} S \\ i \end{pmatrix}\theta^{g-1}\exp\left\{-\theta\left[\frac{T}{K}\left(S'+1\right) + h\right]\right\}$$

Où

$$K = \sum_{i=0}^{S}(-1)^i\begin{pmatrix} S \\ i \end{pmatrix}\left\{\frac{T}{K}\left(S'+1\right) + h\right\}^{-g}$$

Prédiction de$Y = Y(r), 1 \leq r \leq N$

Considérons un échantillon futur de taille$N$, indépendant de l'échantillon sur lequel a porté l'observation. On s'intéresse alors dans cette section à la prédiction de la r^{iéme} statistique d'ordre en utilisant une distribution à priori sur θ de type $G\left(g, h\right)$, et une fonction de perte quadratique.

La densité de Y est :

$$f\left(y/\theta\right) = r\begin{pmatrix} N \\ r \end{pmatrix}\sum_{j=0}^{r-1}(-1)^j\begin{pmatrix} r-1 \\ j \end{pmatrix}\theta\exp\left\{-\theta y(N-r+1-j)\right\}$$

Densité prédictive par par rapport à une fonction de perte quadratique est :

$$p\left(y/x\right) = \int_0^{+\infty} f\left(y/\theta\right)p\left(\theta/x\right)d\theta$$

$$= \frac{r\binom{N}{r}}{k\Gamma\left(g\right)}\sum_{i=0}^{S}\sum_{j=0}^{r-1}(-1)^{i+j}\binom{S}{i}\binom{r-1}{j}\int_0^{+\infty}\theta^g\exp\left\{-\theta\left[(N-r+1-j) + H + \frac{T}{k}i\right]\right\}d\theta$$

$$= \frac{gr\binom{N}{r}}{k}\sum_{i=0}^{S}\sum_{j=0}^{r-1}(-1)^{i+j}\binom{S}{i}\binom{r-1}{j}\left\{(N-r+1-j)\,g+H+\frac{T}{k}i\right\}^{-(g+1)}$$

où

$$H = \frac{T}{k}S' + h$$

la fonction de répartition prédictive est alors :

$$F(z/\underline{x}) = \int_0^z f(y/x)\,dy$$

$$\begin{aligned} = \; & \frac{r\binom{N}{r}}{k}\sum_{i=0}^{S}\sum_{j=0}^{r-1}(-1)^{i+j}\binom{S}{i}\binom{r-1}{j}(N-r+1-j)^{-1} \\ & \times\left\{(H+iT/k)^{-g}-\left[z(N-r+1+j)+(H+iT/k)^{-g}\right]\right\} \end{aligned}$$

Sous l'hypothèse $g > 1$;un prédicteur de $Y_{(r)}$ serait :

$$E(Y/\underline{x}) = \frac{r\binom{N}{r}}{(g-1)\,k}\sum_{i=0}^{S}(-1)\begin{pmatrix} S \\ i \end{pmatrix}\{H+iT/k\}^{-(g+1)}\sum_{j=0}^{r-1}(-1)^{j}\binom{r-1}{j}(N-r+1+j)^{-2}$$

cas particulier : $r = 1$

Il s'agit de la prédiction de la $1^{ére}$statistique d'ordre, ce qui se traduit dans un cadre fiabiliste par l'apparition de la 1ére panne. La fonction de densité prédictive et la fonction de répartition prédictive sont respectivement données par

$$f(y/\underline{x}) = \frac{gN}{K}\sum_{i=0}^{S}(-1)^{i}\binom{S}{i}\left\{Ny+H+i\frac{T}{K}\right\}^{-(g+1)}$$

$$F(z/x) = 1-\frac{1}{k}\sum_{i=0}^{S}(-1)^{i}\binom{S}{i}\left[z(N-r+1+j)+H+iT/k\right]^{-g}$$

2.7.2 Prediction de $Y = Y_{(k)}$ quand on dispose d'un test de survie avec renouvellement

On observe durant un temps T, n items dont les durées de vie sont issues de (1). On suppose de plus que tout item tombé en panne dans$[0, T]$est renouvelé et que le renouvellement est instantané.

L'information disponible à la fin du test est le nombre X de renouvellements ayant eu lieu dans $[0, T]$: On s'intéresse à la prédiction de (2.21) et(2.22).

Prediction de $Y = Y_{(k)}; 1 \leq k \leq N$

Les intervalles entre deux pannes consécutives suivent une loi $expo$: Le modèle considéré est donc une superposition de processus de Poisson de paramètre θ On en déduit que X suit une loi de Poisson. Soit $L\left(x/\theta\right)$ la vraisemblance :

$$L\left(x/\theta\right) = \frac{\left(\theta n T\right)^{x}}{x!} \exp\left\{-\theta n T\right\};\ \theta > 0; x = 0, 1, ..$$

Il est clair que la famille de distribution $G\left(g, h\right)$ est une famille de conjuguées naturelle pour θ.Soit $p\left(\theta\right)$ la densité a priori sur θ :

$$p\left(\theta\right) = \frac{h^{g}}{\Gamma\left(g\right)} \theta^{g-1} \exp\left\{-\theta h\right\};\ \ h, g > 0$$

La densité a posteriori sur θ est alors une loi $G\left(G, H\right)$

$$p\left(\theta/x\right) = \frac{H^{G}}{\Gamma\left(G\right)} \theta^{G-1} \exp\left\{-\theta H\right\};$$

où : $H = k + nT$ et : $G = g + x$

Densité prédictive de $Y_{(k)}$

$$p\left(y/x\right) = \int_{0}^{+\infty} p\left(y/\theta\right) p\left(\theta/x\right) d\theta$$

où

$$P\left(y/\theta\right) = k\left(\begin{array}{c} N \\ k \end{array} \sum_{i=0}^{k-1} \left(-1\right)^{i} \binom{K-1}{i}\right) \theta \exp\left\{-\theta y\left(N - k + 1 + i\right)\right\}$$

$$P\left(y/x\right) = GH^{G}k\left(\begin{array}{c} N \\ k \end{array}\right) \sum_{i=0}^{k-1} \left(-1\right)^{i} \binom{K-1}{i} \left[H + y\left(N - k + 1 + i\right)\right]^{-(G+1)}$$

Fonction de répartition prédictive $F(z/x)$

$$F(z/x) = \int_0^z p(y/x)\,dy$$

$$GH^G k \left(\begin{array}{c} N \\ k \end{array} \right) \sum_{i=0}^{k-1} (-1)^i \binom{K-1}{i} \int_0^z [H + y(N-k+1+i)]^{-(G+1)}\,dy$$

$$1 - H^G k \left(\begin{array}{c} N \\ k \end{array} \right) \sum_{i=0}^{k-1} (-1)^i \binom{K-1}{i} [H + z(N-k+1+i)]^{-G}$$

En particulier pour $k = 1$

$$F(z_1/x) = 1 - H^G N (H + N_{z1})^{-G}$$

Remark 25 *Une connaissance à priori non informative sur θ correspond à $g, h \longrightarrow 0$ soit :*

$G = x$ et $H = nT$

En particulier pour $Y_{(1)}$;la fonction de répartition serait :

$F(z_1/x) = 1 - N(nT)^x (nT + N_{Z1})^{-x}$

Un intervalle de prédiction pour $Y_{(1)}$ de la forme $(0, z)$et de niveau $1 - y$ est tel que z soit solution de :

$$N \left\{ \frac{nT}{nT + Nz} \right\}^x = y \tag{2.24}$$

Ce qui nous donne ;

$$z = \frac{nT}{N} \left\{ \left(\begin{array}{c} N \\ \gamma \end{array} \right)^{1/x} - 1 \right\} \tag{2.25}$$

Prédicteur de $Y_{(K)}$ sous l'hypothèse $G > 1$

$$E(Y/x) = \frac{H}{G-1} k \left(\begin{array}{c} N \\ \gamma \end{array} \right) \sum_{i=0}^{k-1} (-1)^i \binom{K-1}{i} (N-k+1+i)^{-2}$$

Pour $k = 1$

Prédiction de Z= $\sum_{i=0}^{k} Y_{(i)} + (N-K)\, Y_{(k)}$

Z suit une loi G(k,θ)de densité :

$$f\,(z/\theta) = \frac{\theta^k}{\Gamma\,(k)} z^{k-1} \exp\,(-\theta z) \quad z > 0;\ \theta > 0.$$

Densité prédictive de Z

La densité prédictive $p\,(z/x)$ de Z est :

$p\,(z/x) = \frac{H^G}{\Gamma(k)\Gamma(G)} z^{k-1} = \int\limits_0^\infty \theta^{k+G-1} \exp\left\{-\theta\,(H+z)\right\} d\theta$

$= \frac{H^G}{B(k,G)} \times \frac{Z^{k-1}}{(H+z)^{K+G}}; z \geq 0.$

On voit que la loi de (Z/H) sachant que x suit une loi Béta, de typeII, de paramètres G et H, qu'on peut ramener à une loi de Fisher en faisant la transformation suivante :

$$\frac{G}{KH} Z \to F\,(2k, 2G)$$

Fonction de répartition prédictive $F\,(z/x)$

$$F\,(y/x) = \int\limits_0^z p\,(y/x)\, dy = I_{Z_{H+Z}}\,(k,G)$$

Un intervalle de prédiction de$Z$ de la forme $(0,Z)$ et de niveau $(1-\gamma)$ est tel que z soit la solution de

$$I_{z/_{H+Z}}\,(k,G) = (1-\gamma) \tag{2.26}$$

Prédicteur de Z par rapport à une fonction de perte quadratique et sous l'hypothèse $G > 1$

$$\begin{aligned} E\,(z/x) &= \frac{H^G}{B\,(k,G)} \int\limits_0^\infty \frac{Z^{k-1}}{(z+H)^{K+G}} dz \\ &= H \frac{k}{(G-1)} \end{aligned}$$

Remark 26 *pour une connaissance à priori non informative sur θcorrespondant à* $g, h \to 0$

$$E\left(Y_1/x\right) = \frac{H}{\left(G-1\right)N^2} \tag{2.27}$$

un prédicteur de Zserait, sous l'hypothèse $x > 1$.

$$E\left(Z/x\right) = K\frac{nT}{x-1} \tag{2.28}$$

CONCLUSION

On obtient des prédicateurs de (2.21) et (2.22) de forme analytique très simple et donc facilement exploitables, et ce malgré l'utilisation de plans d'expérience nettement moins contraignants et onéreux que les plans d'expériences complets ou les plans d'expériences avec censures à droite.

Chapitre 3

Modèle de Rayleigh

La distribution de Rayleigh est fortement utilisée dans les modèles de survie, en particulier dans des études cliniques, car elle a la particularité d'avoir un taux de panne linéaire par rapport au temps $h(t) = \frac{t}{\sigma 2}$. Plusieurs auteurs se sont intéréssés à ce modèle ; Howlader H.A. et Hossain A.(1995) ont étudié le problème de l'estimation du paramètre et de la fonction de fiabilité avec des données censurées et une fonction de perte quadratique ; Abdelffatah, A.M., Hassan, Amel S. et Zieden, D.M. (2006) ont étudié l'efficacité des estimateurs obtenus par une approche classique du maximum de vraisemblance sous différentes sortes de données.

La fonction de densité $f(x)$ et la fonction de fiabilité $(S(t))$ et le taux de panne $h(t)$ dans un modèle de Rayleigh sont données par :

$$f(x) = \frac{x}{\sigma^2} \exp\left(-\frac{x^2}{2\sigma^2}\right) \; ; \; x > 0, \sigma > 0 \tag{3.1}$$

$$S(t) = 1 - F(t) = \exp\left(-\frac{t^2}{2\sigma^2}\right) ; \; t > 0 \tag{3.2}$$

et

$$h(t) = \frac{t}{\sigma^2} \tag{3.3}$$

On se propose dans cet article d'étudier les estimateurs Bayesiens du paramètre, de la fonction de fiabilité et du taux de panne sous une fonction de perte quadratique et sous une fonction de perte asymétrique en présence d'un échantillon censuré de type II et en considérant une loi a priori non informative (section 2), puis une loi a priori conjuguée naturelle (section 3). La section quatre est consacrée à une étude comparative à l'aide du critère de Pittman.

3.1 Estimation Bayesienne avec une loi a priori non informative

Soit $\left(X_{(1)}, X_{(2)}, ..., X_{(r)}, ..., X_{(n)}\right)$ un échantillon de taille n censuré en $X_{(r)}$, la vraisemblance s'écrit :

$$L\left(x,\sigma\right) \propto \frac{1}{\sigma^{2r}} \exp\left(-\frac{1}{2\sigma^2} T_r\right)$$

où :

$$T_r = \sum_{i=1}^{r} x_i^2 + (n-r)x_r^2$$

Il est aisé de montrer que la loi a priori définie par Jeffrey est

$$\pi_1(\sigma) = \left|I_1\left(\sigma\right)\right|^{\frac{1}{2}} = \left|-E\frac{\partial^2 \ln f}{\partial \sigma^2}\right|^{\frac{1}{2}} \propto \frac{1}{\sigma}$$

La loi a postériori est alors :

$$\pi_1(\sigma/x) = \frac{L\left(x,\sigma\right)\pi_1(\sigma)}{\int\limits_0^\infty L\left(x,\sigma\right)\pi_1(\sigma)d\sigma} = \frac{(T_r)^r}{\Gamma(r)} \times \frac{1}{2^{r-1}} \times \sigma^{-2r-1} \times \exp(-\frac{T_r}{2\sigma^2}) \quad (3.4)$$

Où x représente le vecteur des observations.

3.1.1 fonctions de pertes

fonction de perte quadratique

Soit $\overset{\curlywedge}{\theta}$ un estimateur de θ, une fonction de perte quadratique est définie par $L_1(\theta, \overset{\curlywedge}{\theta}) = (\theta - \overset{\curlywedge}{\theta})^2$ a été proposée par Legendre (1805) et Gauss (1810), elle est trés utilisée en litérature. L'estimateur Bayesien de θ est alors égal à sa moyenne a posteriori, soit $\overset{\curlywedge}{\theta}_B = E(\theta/x)$.

fonction de perte LINEX

Considérons la fonction de perte convexe suivante

$$L(\Delta) \propto e^{a\Delta} - a\Delta - 1, a \neq 0 \quad (3.5)$$

Le signe de a et sa grandeur absolue représentent respectivement la direction et le degré d'asymétrie. Pour $a \to 0$, on retrouve la fonction de perte quadratique. Varian (1975) a considéré la fonction de perte $(3,5)$ pour $\Delta_1 = \hat{\theta} - \theta$; la fonction $L(\Delta_1)$ est alors appelée fonction de perte LINEX. Soit $E_p\left(L(\Delta_1)\right)$ l'espérance a posteriori de $L(\Delta_1)$, l'estimateur Bayesien de θ sous cette fonction de perte est noté θ_{LB} il correspond à la valeur de $\hat{\theta}$ qui minimise $E_p\left(L(\Delta_1)\right)$.

$$E_p\left(L\left(\Delta_1\right)\right) = \exp\left(a\hat{\theta}\right) E_p\left(\exp -a\theta\right) + aE\left(\theta\right) - a\hat{\theta} - 1$$

On dérive l'expression ci-dessus par rapport à $\hat{\theta}$ et on égale à zéro

$$\frac{\partial E_p\left(L\left(\Delta_1\right)\right)}{\partial\hat{\theta}} = a\exp\left(a\hat{\theta}\right) E_p\left(\exp\left(-a\theta\right)\right) - a = 0 \tag{3.6}$$

La solution de l'équation $(3,6)$ est

$$\hat{\theta}_{LB} = -\frac{1}{a}\ln(E(e^{-a\theta})) \tag{3.7}$$

Considérons la la fonction de perte $L\left(\Delta_2\right)$; où $\Delta_2 = \left(\frac{\hat{\theta}}{\theta}\right)^2 - 1$; cette fonction de perte a été utilisée par plusieurs auteurs dont Zellner (2006) et (2009). On minimise l'espérance a posteriori $E_p\left(L\left(\Delta_2\right)\right)$:

$$\begin{aligned} E_p\left(L\left(\Delta_2\right)\right) &= E_p\left[\exp\left(\left(\frac{\hat{\theta}}{\theta}\right)^2 - 1\right) - a\left(\left(\frac{\hat{\theta}}{\theta}\right)^2 - 1\right) - 1\right] \\ &= \exp\left(-a\right) E_p\left(\exp\left(\frac{\hat{\theta}}{\theta}\right)^2\right) - a\left(\frac{\hat{\theta}}{\theta}\right)^2 - 1 \end{aligned}$$

En dérivant $E_p\left(L\left(\Delta_2\right)\right)$ par rapport à $\hat{\theta}$ et en égalant à zéro, on obtient l'estimateur Bayesien $\hat{\theta}_{LB}$ de θ sous la fonction de perte $L\left(\Delta_2\right)$

$$\frac{\partial E_p\left(L\left(\Delta_2\right)\right)}{\partial\hat{\theta}} = 2a\exp\left(-a\right) E_p\left(\frac{\hat{\theta}}{\theta^2}\exp\left(a\left(\frac{\hat{\theta}}{\theta}\right)^2\right)\right) - 2aE_p\left(\frac{\hat{\theta}}{\theta^2}\right) = 0$$

On obtient alors $\overset{\curlywedge}{\theta}_{LB}$ comme solution de l'équation

$$E_p\left[\frac{1}{\theta^2}\exp\left(a\left(\frac{\overset{\curlywedge}{\theta}{}^2_{LB}}{\theta^2}\right)\right)\right]=\exp\left(a\right)E_p\left(\frac{1}{\theta^2}\right) \tag{3.8}$$

3.1.2 Estimation du paramètre σ

L'estimateur du maximum de vraisemblance noté $\overset{\curlywedge}{\sigma}_{MV}$ est obtenu en résolvant l'équation $\dfrac{\partial \ln L\left(x,\sigma\right)}{\partial\sigma}=0$; on obtient :

$$\overset{\curlywedge}{\sigma}_{MV}=\sqrt{\frac{T_r}{2r}} \tag{3.9}$$

Par rapport à une fonction de perte quadratique, et avec une loi a priori vague sur σ; l'estimateur de σ noté $\sigma_{v,Q}$ est obtenu en calculant son espérance par rapport à la densité a postériori :

$$\sigma_{v,Q}=\int_0^\infty \sigma\pi_1(\sigma/x)d\sigma=\frac{\left(T_r\right)^r}{\Gamma(r)}\times\frac{1}{2^{r-1}}\int_0^\infty\sigma^{-2r}\times\exp(-\frac{T_r}{2\sigma^2})d\sigma=\frac{\Gamma\left(r-\frac{1}{2}\right)}{\Gamma\left(r\right)}\left(\frac{T_r}{2}\right)^{\frac{1}{2}} \tag{3.10}$$

Par rapport à la fonction de perte asymétrique $L\left(\Delta_2\right)$ et avec une loi a priori vague, un estimateur de σ noté $\sigma_{v,L}$ est la solution de l'équation donnée en $(3,8)$ où :

$$E_p\left[\frac{\sigma_{v,L}}{\sigma^2}\exp\left(a\left(\frac{\sigma^2_{v,L}}{\theta^2}\right)\right)\right]=\frac{\left(T_r\right)^r}{\Gamma(r)}\frac{\sigma_{v,L}}{2^{r-1}}\int_0^\infty\frac{1}{\sigma^{2r+3}}\exp\left(-\tfrac{1}{2\sigma^2}\left(T_r-2a\sigma^2_{v,L}\right)\right)d\sigma$$

et

$$\exp\left(a\right)E_p\left(\frac{\sigma_{v,L}}{\sigma^2}\right)=\frac{\left(T_r\right)^r}{\Gamma(r)}\times\frac{\sigma_{v,L}}{2^{r-1}}\int_0^\infty\frac{1}{\sigma^{2r+3}}\exp\left(-\tfrac{1}{2\sigma^2}\left(T_r\right)\right)d\sigma$$

Aprés quelques manipulations algébrique, on obtient :

$$\sigma_{v,L}=\left[\frac{T_r}{2a}\left(1-\exp\left(-\frac{a}{r+1}\right)\right)\right]^{\dfrac{1}{2}} \tag{3.11}$$

3.1.3 Estimation de la fonction de fiabilité

Pour avoir l'estimateur $S_{MV}\left(t\right)$ de la fonction de fiabilité $S\left(t\right)$, il suffit de remplacer dans l'expression de $S\left(t\right)$ donnée en $(3,2)$ σ par $\overset{\curlywedge}{\sigma}_{MV}$donnée en $(3,9)$:

$$S_{MV}\left(t\right)=\exp(-\frac{rt^2}{T_r})$$

L'estimateur Bayesien de $S(t)$ par rapport à une fonction de perte quadratique et une loi a priori vague est $S_{Q,v}(t)$:

$$S_{Q,v}(t) = \int_0^\infty S(t)\,\pi_1(\sigma/x)d\sigma = \left(\frac{T_r}{T_r + t^2}\right)^r \tag{3.12}$$

L'estimateur Bayesien par rapport à une fonction de perte Linex est noté $S_{L,v}(t)$, pour le calculer, on fait le changement de variable suivant :

$S(t) = \exp\left(-\frac{t^2}{2\sigma^2}\right) = \gamma \Longrightarrow \sigma = \left(-\dfrac{t^2}{2\ln\gamma}\right)^{\frac{1}{2}}$; on reécrit la densité a posteriori donnée en $(3,4)$ en fonction de γ.

$$\begin{aligned}
\pi_1(\gamma/x) &= \frac{(T_r)^r}{\Gamma(r)} \times \frac{1}{2^{r-1}}\frac{1}{2}\left(-\frac{t^2}{2\ln\gamma}\right)^{\frac{1}{2}}\left(\frac{t^2}{2\gamma(\ln\gamma)^2}\right)\left(-\frac{t^2}{2\ln\gamma}\right)^{-r-\frac{1}{2}}(\gamma)^{\frac{T_r}{t^2}} \\
&= \left(\frac{Tr}{t^2}\right)^r \frac{1}{\Gamma(r)}(\gamma)^{\frac{Tr}{t^2}-1}(-\ln\gamma)^{r-1} \qquad 0 \le \gamma \le 1
\end{aligned}$$

En utilisant la fonction de perte $L(\Delta_1)$; un estimateur Bayesien $\gamma_{L,v}$ de γ d'aprés $(2,3)$:

$$\begin{aligned}
\gamma_{L,v} &= -\frac{1}{a}\ln E_p(\exp(-a\gamma)) \qquad (3.13) \\
&= -\frac{1}{a}\ln\left[\left(\frac{Tr}{t^2}\right)^r \frac{1}{\Gamma(r)}\int_0^1 \exp(-a\gamma)(\gamma)^{\frac{Tr}{t^2}-1}(-\ln\gamma)^{r-1}d\gamma\right] \\
&= -\frac{1}{a}\ln\left[\sum_{j=0}^{k}\frac{(-a)^j}{j!}\left(1 + j\frac{t^2}{Tr}\right)^{-r}\right]
\end{aligned}$$

Ce dernier résultat est obtenu en utilsant un développement de $(\exp(-a\gamma))$ à l'ordre k dans un voisinage de zéro et en faisant le changement de variable $u = (-\ln\gamma)$ pour le calcul de l'intégrale.

3.1.4 Estimation du taux de panne

Lestimateur du maximum de vraisemblance de $h(t)$ noté $h_{MV}(t)$ est simplement obtenu en remplaçant σ par σ_{MV} dans l'expression de $h(t)$, ce qui nous donne :

$$h_{MV}(t) = 2r\frac{t}{T_r} \tag{3.14}$$

Par rapport à une fonction de perte quadratique et une connaissance vague sur σ, l'estimateur Bayesien de $h(t)$ noté $h_{v,Q}(t)$ est :

$$\begin{aligned} h_{v,Q}(t) &= \int_0^{\infty} h(t)\,\pi_1(\sigma/x)d\sigma = \frac{(T_r)^r}{\Gamma(r)} \times \frac{t}{2^{r-1}} \int_0^{\infty} \sigma^{-2r-3} \times \exp(-\frac{T_r}{2\sigma^2})d\sigma \\ h_{v,Q}(t) &= 2r\frac{t}{T_r} \end{aligned} \tag{3.15}$$

Remarque : Les estimateurs de $h(t)$ obtenus par maximum de vraisemblance et par une approche Bayesienne avec une loi a priori non informative et une fonction de perte quadratique sont identiques.

Les fonctions de perte asymétriques $L(\Delta_1)$ et $L(\Delta_2)$ ne sont pas appropriées pour avoir une forme analytique simple d'un estimateur Bayesien de $h(t)$; c'est pourquoi, on définit $\Delta = \left(\frac{\theta}{\hat{\theta}} - 1\right)$ on remplace dans la fonction de perte donnée en $(3,5)$, puis on prend l'espérance a postériori qu'on dérive et qu'on égale à zéro pour trouver la valeur de θ notée θ_{LB} qui la minimise.

$$E_p(L(\Delta)) = E_p\left(\exp a\left(\frac{\theta}{\hat{\theta}} - 1\right)\right) - aE_p\left(\frac{\theta}{\hat{\theta}} - 1\right) - 1$$

$$\frac{\partial E_p(L(\Delta))}{\partial\hat{\theta}} = \exp(-a)\,E_p\left(-a\frac{\theta}{\hat{\theta}^2}\exp\left(\frac{\theta}{\hat{\theta}}\right)\right) + aE_p\left(\frac{\theta}{\hat{\theta}^2}\right) = 0$$

L'estimateur Bayesien θ_{LB} par rapport à la fonction de perte $L(\Delta)$ est alors solution de l'équation :

$$\exp(-a)\,E_p\left(\theta\exp a\left(\frac{\theta}{\hat{\theta}_{LB}}\right)\right) = E_p(\theta) \tag{3.16}$$

On pose $h(t) = \theta$, soit $\frac{t}{\sigma^2} = \theta \Longrightarrow \sigma = \left(\frac{t}{\theta}\right)^{\frac{1}{2}}$; on reécrit la densité a postériori $(3,4)$ en fonction de θ :

$$\begin{aligned} \pi_1(\theta/x) &= \frac{(T_r)^r}{\Gamma(r)} \times \frac{1}{2^{r-1}}\left(\frac{t}{\theta}\right)^{-r-\frac{1}{2}}\exp\left(-\frac{T_r}{2t}\theta\right)\frac{1}{2}\left(\frac{t}{\theta}\right)^{-\frac{1}{2}}\left(\frac{t}{\theta^2}\right) \\ &= \left(\frac{T_r}{2t}\right)^r \frac{1}{\Gamma(r)}\theta^{r-1}\exp\left(-\frac{T_r}{2t}\theta\right) \qquad \theta \geq 0 \end{aligned}$$

On remarque que la loi de θ a postériori est celle d'une loi gamma $\mathcal{G}\left(r, \dfrac{T_r}{2t}\right)$.

On résoud l'équation $(3,16)$, où $\overset{\curlywedge}{\theta}_{LB}$ représente l'estimateur du taux de panne $h\left(t\right)$ et qu'on note par souci d'homogeneité d'écriture $h_{v,L}\left(t\right)$:

$$\exp\left(-a\right)\left(\frac{T_r}{2t}\right)^r \frac{1}{\left[\dfrac{T_r}{2t}-\dfrac{a}{h_{v,L}\left(t\right)}\right]}=\frac{2t}{T_r}$$

$$h_{v,L}\left(t\right)=a\frac{2t}{T_r}\left[1-\exp\left(-\frac{a}{r+1}\right)\right]^{-1} \tag{3.17}$$

3.2 Estimation Bayesienne avec une loi a priori conjuguée naturelle

La loi a priori conjugué naturelle est définie comme suit :

$$\pi_2\left(\sigma\right)\propto\frac{1}{\sigma^{\alpha+1}}\exp(-\frac{\beta}{2\sigma^2});\ \alpha,\beta>0$$

La loi a postériori est alors :

$$\pi_2(\sigma/x)=\frac{\left(T_r+\beta\right)^{r+\frac{\alpha}{2}}}{2^{r+\frac{\alpha}{2}-1}\Gamma\left(r+\frac{\alpha}{2}\right)}\times\frac{1}{\sigma^{2r+\alpha+1}}\exp(-\frac{1}{2\sigma^2}(T_r+\beta)) \tag{3.18}$$

3.2.1 Estimation du paramètre σ :

Toujours par rapport à une fonction de perte quadratique, mais avec une loi a priori conjugué naturelle sur σ; l'estimateur de σ noté $\sigma_{B,Q}$ est obtenu en calculant son espérance par rapport à la densité a postériori :

$$\sigma_{B,Q}=\int_0^\infty\sigma\pi_2(\sigma/x)d\sigma=\sqrt{\frac{T_r+\beta}{2}}\times\frac{\Gamma\left(r+\frac{\alpha}{2}-\frac{1}{2}\right)}{\Gamma\left(r+\frac{\alpha}{2}\right)} \tag{3.19}$$

l'estimation Bayesienne de σ par rapport à la fonction de perte Δ_2 et une loi a priori conjuguée naturelle sur σ ,un estimateur de σ noté $\sigma_{B,L}$ est la solution de l'équation :

$$E_p\left[\frac{\sigma_{B,L}}{\sigma^2}\exp\left(a\left(\frac{\sigma_{B,L}^2}{\theta^2}\right)\right)\right]=2^{r+\frac{\alpha}{2}}\times\hat{\sigma}_{BL}\left[\frac{\Gamma\left(r+\frac{\alpha}{2}+1\right)}{\left(T_r+\beta-2a\hat{\sigma}^2_{LB}\right)^{r+\frac{\alpha}{2}+1}}\right]$$

$$\exp\left(a\right)E_p\left(\frac{\sigma_{B,L}}{\sigma^2}\right)=\hat{\sigma}_{BL}e^a\times\frac{\Gamma\left(r+\frac{\alpha}{2}+1\right)}{\left(T_r+\beta\right)^{r+\frac{\alpha}{2}+1}}$$

$$\hat{\sigma}_{BL} = \left[\frac{(T_r+\beta)}{2a} \left(1 - \exp(-\frac{a}{r+\frac{\alpha}{2}+1} \right) \right]^{\frac{1}{2}} \tag{3.20}$$

3.2.2 Estimation de la fonction de fiabilité :

L'estimateur Bayesien de $S(t)$ par rapport à une fonction de perte quadratique et une loi a priori conjugué naturelle est $S_{Q,B}(t)$:

$$S_{Q,B}(t) = \int_0^\infty S(t)\,\pi_2(\sigma/x)d\sigma = (\frac{T_r+\beta}{T_r+\beta+t^2})^{r+\frac{\alpha}{2}}$$

$$S_{Q,B}(t) = (\frac{T_r+\beta}{T_r+\beta+t^2})^{r+\frac{\alpha}{2}} \tag{3.21}$$

avec une loi conjugué naturelle, l'estimateur Bayesien par rapport à une fonction de perte Linex est noté $S_{L,B}(t)$, pour le calculer, on fait le changement de variable suivant :

$S(t) = \exp\left(-\frac{t^2}{2\sigma^2}\right) = \gamma \Longrightarrow \sigma = \left(-\frac{t^2}{2\ln\gamma}\right)^{\frac{1}{2}}$; on reécrit la densité a posteriori donnée en $(3,18)$ en fonction de γ.

$$\pi_2(\gamma/x) = \frac{(T_r+\beta)^{r+\frac{\alpha}{2}}}{\Gamma\left(r+\frac{\alpha}{2}\right)} \times \frac{1}{(t^2)^{r+\frac{\alpha}{2}}} \times (\gamma)^{\frac{T_r+\beta}{t^2}-1} \times (-\ln\gamma)^{r+\frac{\alpha}{2}-1} \tag{3.22}$$

En utilisant la fonction de perte $L(\Delta_1)$; un estimateur Bayesien $\gamma_{L,B}$ serait :

$$\begin{aligned}
\gamma_{L,B} &= -\frac{1}{a}\ln E_p\left(\exp(-a\gamma)\right) \qquad (3.23)\\
&= -\frac{1}{a}\ln\left[\frac{(T_r+\beta)^{r+\frac{\alpha}{2}}}{\Gamma\left(r+\frac{\alpha}{2}\right)} \times \frac{1}{(t^2)^{r+\frac{\alpha}{2}}} \int_0^1 \exp(-a\gamma)\,(\gamma)^{\dfrac{Tr+\beta}{t^2}-1}(-\ln\gamma)^{r+\frac{\alpha}{2}-1}d\gamma\right]\\
&= -\frac{1}{a}\ln\left[\sum_{j=0}^{k}\frac{(-a)^j}{j!}\left(1+\frac{jt^2}{T_r+\beta}\right)^{\left(-r-\frac{\alpha}{2}\right)}\right]
\end{aligned}$$

3.2.3 estimation de taux de panne :

Par rapport à une fonction de perte quadratique

$$h_{Q,B} = \frac{2\left(r+\frac{\alpha}{2}\right)t}{T_r+\beta} \tag{3.24}$$

Par rapport à une fonction de cout Linex :

On remarque que la loi de θ a postériori est celle d'une loi gamma $\mathcal{G}\left(r+\frac{\alpha}{2}, \frac{T_r+\beta}{2t}\right)$.

On résoud cette équation où $\overset{\curlywedge}{\theta}_{LB}$ représente l'estimateur du taux de panne $h(t)$ et qu'on note $h_{B,L}(t)$:

$$\exp(-a)\left(\frac{T_r+\beta}{2t}\right)^{r+\frac{\alpha}{2}} \frac{1}{\left[\frac{T_r+\beta}{2t}-\frac{a}{h_{B,L}(t)}\right]^{r+\frac{\alpha}{2}+1}} = \frac{2t}{T_r+\beta}$$

:

$$h_{B,L}(t) = a\frac{2t}{T_r+\beta}\left[1-\exp\left(-\frac{a}{r+\frac{\alpha}{2}+1}\right)\right]^{-1} \tag{3.25}$$

3.3 Simulation :

1) On se donne des valeurs de α et β, on genére une loi définie en (3.1) ; on en déduit une valeur pour σ.

2) On génére $N = 10000$ échantillons de taille n d'une loi de Rayleigh de paramètre σ, on prend trois valeurs de n $(n = 20, 30, 50)$ et deux valeurs pour r, pour avoir des taux de censure croissants.

3) On calcule les différents estimateurs de σ, $S(t)$, et $h(t)$; par maximum de vraisemblance (noté MV), avec une approche Bayesienne sous une fonction de perte quadratique (noté MQ) et enfin avec une fonction de perte asymétrique (noté Li) pour quatre valeurs du coefficient d'asymétrie a $(a = -1; -0.5; 0.5; 1)$.

4) Pour chaque estimateur, on calcule l'erreur quadratique moyenne notée (MSE) ; ϕ pouvant être σ, $S(t)$, $h(t)$ et $\overset{\curlywedge}{\phi}$ leur estimateur respectivement.

$$MSE = \frac{1}{10000}\sum_{i=1}^{10000}\left(\phi-\overset{\curlywedge}{\phi}\right)^2$$

loi a priori vague :

Table 1 :Estimation de $\sigma, S(t)$ et $h(t)$ lorsque $(\alpha = 0, \beta = 0, \sigma = \Gamma(1/2)$, $t = 0.75 \quad S(0.75) = 0.8360, h(0.75) = 0.2387)$

(n,r)		$_{MV}$ (MSE)	$_{MQ}$ (MSE)	$_{Li}$ (MSE)			
				$a=-0.5$	$a=-1$	$a=0.5$	$a=1$
	σ	1.7494	1.8187	1.6872	1.7084	1.6489	1.6328
		(0.0005)	(0.0021)	(0.0072)	(0.0041)	(0.0152)	(0,0194)
$(20,10)$	$S(t)$	0.9058	0.9063	0.8841	0.8402	0.8626	0.6827
		(0.0048)	(0.0049)	(0.0023)	(0.00017)	(0.0007)	(0.0235)
	$h(t)$	0.2651	0.2651	0.2850	0.2792	0.2982	0.3043
		(0.0006)	(0.0006)	(0.0021)	(0.0016)	(0.0035)	(0.0043)
	σ	1.7513	1.8206	1.6890	1.7081	1.6512	1.6308
		(0.00044)	(0.0023)	(0.0069)	(0.0041)	(0.0146)	(0.0200)
$(30,10)$	$S(t)$	0.9059	0.9063	0.8841	0.8403	0.8625	0.6826
		(0.00487)	(0.0049)	(0.0023)	(0.00002)	(0.0007)	(0.0235)
	$h(t)$	0.2652	0.2652	0.2851	0.2788	0.2984	0.3057
		(0.0007)	(0.0007)	(0.0021)	(0.00160)	(0.00357)	(0.0044)
	σ	1.7478	1.8170	1.6856	1.7090	1.6506	1.6344
		(0.0006)	(0.0019)	(0.00753)	(0.00401)	(0.0148)	(0.0190)
$(50,10)$	$S(t)$	0.9057	0.9062	0.8840	0.8405	0.8626	0.6827
		(0.0048)	(0.0049)	(0.0022)	(0.00002)	(0.0007)	(0.0235)
	$h(t)$	0.2656	0.2656	0.2856	0.2780	0.2982	0.3043
		(0.0007)	(0.0007)	(0.0021)	(0.0015)	(0.00354)	(0.0043)

	σ	1.7557	1.8012	1.7157	1.7285	1.6856	1.6738
		(0.0002)	(0.0008)	(0.0032)	(0.00193)	(0.00752)	(0.0231)
$(20,15)$	$S(t)$	0.9086	0.9089	0.88675	0.8426	0.8649	0.6838
		(0.0052)	(0.0053)	(0.0025)	$(4e-05)$	(0.0008)	(0.0231)
	$h(t)$	0.2565	0.2565	0.2684	0.2645	0.2779	0.2822
		(0.0003)	(0.0003)	(0.0008)	(0.00066)	(0.0015)	(0.0018)
	σ	1.7625	1.8082	1.7200	1.7298	1.6901	1.6755
		$(9e-05)$	(0.0012)	(0.0027)	(0.0018)	(0.0067)	(0.0093)
$(30,15)$	$S(t)$	0.9093	0.9096	0.8871	0.8426	0.8651	0.6838
		(0.0053)	(0.0054)	(0.0026)	$(4e-05)$	(0.0008)	(0.0231)
	$h(t)$	0.2543	0.2543	0.2671	0.2642	0.2769	0.2815
		(0.0002)	(0.0002)	(0.0008)	(0.0006)	(0.0014)	(0.0018)
	σ	1.7551	1.8006	1.7119	1.7299	1.6846	1.6732
		(0.0002)	(0.0007)	(0.0036)	(0.0018)	(0.0077)	(0.0098)
$(50,15)$	$S(t)$	0.9085	0.9088	0.8864	0.8427	0.8647	0.6837
		(0.0052)	(0.0052)	(0.0025)	$(4e-05)$	(0.0008)	(0.0231)
	$h(t)$	0.2566	0.2566	0.2693	0.2641	0.2786	0.2824
		(0.0003)	(0.0003)	(0.0009)	(0.0006)	(0.0015)	(0.0019)

3.4 Conclusion et Perspectives

Dans cette thèse, on s'est interéssé à deux aspects des problèmes de l'inférence statistique ; à savoir, les problèmes de prédiction de statistiques d'ordre et de fonctions de celles-ci et les problèmes d'estimation à partir d'observations faites sur un échantillon. Les observations les plus utilisées dans la littérature sont, soit complètes ; soit censurées de type II ; nous avons traité le cas où les observations sont groupées par intervallesés de temps préalablement fixés. Ce plan d'expérience a l'avantage d'être moins contraignant d'un point de vue coût. Les résultats obtenus dans le cas d'un modèle exponentiel sont de forme analytique explicite et donc facilement exploitable.

Le deuxième aspect étudié est le problème de l'estimation des paramètres et des caracteristiques de fiabilité dans un modèle de Rayleigh. Ce modèle est largement utilisé en analyse de survie, notament en épidémiologie où les applications sont nombreuses.L'approche utilisée est une approche Bayesienne avec une loi a priori non informative dans un premier temps puis, nous avons considéré le cas d'une loi a priori conjuguée naturelle . En rappelant les résultats obtenus par l'approche classique du maximum de vraisemblance, nous avons mené une étude comparative des différents estimateurs obtenus avec d'une part une fonction de perte quadratique et d'autre part, avec une fonction de perte asymétrique LINEX.

Une étude par simulation sur 10000 échantillons de différentes tailles $(n = 20, 30, 50)$ d'une loi de Rayleigh avec différents taux de censure a conduit aux conclusions suivantes :

1) Les erreurs quadratiques des estimateurs du paramètre, de la fonction de fiabilité et du taux de panne diminuent quand le taux de censue augmente. Cependant, elles restent relativement petites dans l'ensemble.

2) Les erreurs quadratiques les plus petites pour toutes les tailles d'échantillon et pour tout taux de censure correspondent aux estimateurs obtenus par une approche Bayesienne avec une fonction de perte LINEX et un coefficient d'asymétrie égal à $a = -1$.

Il existe plusieurs axe de développement de ces deux aspects traités dans cette thèse, le premier consiste à utiliser des données progressivement censurées pour les deux modèles étudiés, à savoir le modèle exponentiel et le modèle de Rayleigh. Une autre approche consiste en l'utilisation des modèles accélérés.

Bibliographie

[1] Arnold BC.(1989). Bayesian estimation and prediction for Pareto data.J.A.S.A., Vol 84, N°408 1079-1084.

[2] Bancroft G.A. and Dunsmore I.R. (1976)."Prédictive distribution in life testes under competing causes of failure".-Biometrica, Vol 63, N°1 195-217

[3] Bratcher-Schucany et Hunt.(1971). Bayesian prédiction and population size assumptions. Technometrics, Vol 13, 678-684.

[4] Congdon Peter (2006). Bayesian Statistical Modelling Second edition. Wiley Eds

[5] Bacha, M. (1996). Inférence statistique pour des modèles de durées de vie et applications, Thése de doctorat, Université de Rouen.

[6] Bacha, M. & Celeux, G.(1996). BRM-IS : Un algorithme d'estimation bayésienne pour modèles à données incomplètes, Proceeding des XXVIII journées de statistique, Québec.

[7] Bacha, M. ,Celeux, G, Idéé, E., Lannoy, A & Vasseur, D.(1998). Esitimation de modèles de durées de vie fortement censurées, Eyrolles.

[8] Bain, L. & Engelhardt, M. (1991). Statistical Analysis of Reliability and Life Testing Models, Marcel Dekker.

[9] Basu, S, Sen, A. & M. Banerjee. (2003). Bayesian analysis of competing risks with partially masked cause of failure, A pplied Statistics, 52, pp. 77-93.

[10] Berger, J.O. (1985). Statistical Decision Theory and Bayesian Analysis (2 nd édition), Springer-Verlag.

[11] Berger, J.O. & Bernardo, J.M.(1992). On the development of reference priors (with discusion).

[12] Berger, J.O. & Sun, D.(1993). Bayesian analysis for the Poly- Weibull Distribution, JASA, 88, pp. 1412-1418.

[13] Bernardo, J.M. (1997). Reference Posterior Distributions for Bayesian Inférence, J.R. Statis.Soc, 41, pp. 113-147

[14] Block,H.W. and Basu,A.P. (1974) : A continuous Bivariate exponential extension. J.Amer.Statist.Assoc., Vol 69, 1031-1037.

[15] Cappé, O., Guillin, A., Marin, J., M. & Robert, C.P.(2004). Population Monte Carlo, Journal of Computational and Graphical Statistics, 13, pp. 907-929.

[16] Calabria R., Pulcini G. (1994) : An Engineering Approach to Bayes estimation for the Weibull distribution. Micron Electron Reliab., 34 ; 789-802.

[17] Castanier, B. (1997). Estimation des variances pour des modèles de durées de vie censurées, Mémoire de DEA, INRIA Rhône-Alpes.

[18] Celeux, G. (1996). Estimation of failure times involving Weibul distributions via stochastic algorithms, In : xvii émes Rencontres Franco-Belges de Statisticiens, Marne-la-Vallée.

[19] Chadli, A. (2004).Prédictions Bayésiennes en fiabilité : Cas d'une loi exponentielle, Publications de l'institut de statistiques de Paris. XXXXVIII, fasc. 1-2, 3-17

[20] Chadli, A., Meradji A.(2014)Bayesians predictions of order statistics with grouped data : The case of an exponential law.Rev. Sci. Technol., Synthèse 29 : 00 -00

[21] DeGroot M.H. (1970) : Optimal Statistical Decisions. New-York, McGraw-Hill.

[22] Dunsmore I.R. 1974,The Bayesian predictive distribution in life testing models. Technometrics, Vol 16 N°3 445-450.

[23] Dunsmore I.R. 1974,Asymptotic prediction analysis. Biometrica, vol 63, 627-630.

[24] Erkanli A., Mazzukhi TA.and Soyer R.(1998). Bayesian computation for a class of reliability growth models. Technometrics, Vol 40, $N°1$ pp. 14-23.

[25] Evans I.G. and Nigm A.H.M.1980,Bayesian prediction for the left trancated exponentiel distribution. Technometrics, Vol 22, N°2 ,329-340.

[26] Evans I.G. and Nigm A.H.M.1980, Bayesian 1-sample prediction for the 2-parameter distribution. I.E.E.E. transaction on reliability, Vol R-29 N°5 , 218-229.

[27] Fountain R.L. (2000) A class of closeness criteria. Communication in statistics, theory and methods, 29(8), 1865-1883.

[28] Freund, J.A (1961). bivariate Extension of the exponential distribution. J.Amer.Statist.Assoc. Vol 56, 971-977.

[29] Fuller, W.A (1982). Closeness of estimators. Encyclopedia of statistical sciences, vol 2, Wiley.

[30] Gauss, C.F. (1810). Least Squares method for the Combinations of Observations. Translated by J. Bertrand (1955). Mallet-Bachelier, Paris.

[31] Guillin, A., Marin, J., M. & Robert & C.P. (2005). Estimation bayésienne approximative par échantillonnage préférentiel, Revue de Statistique Appliquée, 54, pp. 79-95.

[32] Hanagal D.D. and Ahmadi K.A. 2009. Bayesian estimation of the parameters of bivariate exponential distribution.Communications in Statistics - Simulation and Computation, 38. 1391-1413.

[33] Hartigan, J.A. (1983). Bayes ′ Théory, New York :