



# Prédictions Bayésiennes en fiabilité : Cas d'une loi exponentielle

A. Chadli

## ► To cite this version:

A. Chadli. Prédictions Bayésiennes en fiabilité : Cas d'une loi exponentielle. Annales de l'ISUP, 2004, XXXXVIII (1-2), pp.3-17. <hal-03640646>

**HAL Id: hal-03640646**

**<https://hal.science/hal-03640646v1>**

Submitted on 13 Apr 2022

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

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



HAL Authorization

# Prédictions Bayésiennes en fiabilité : Cas d'une loi exponentielle

A. Chadli\*

**Résumé :** Dans ce travail, nous nous intéressons à la prédiction Bayésienne de statistiques d'ordre et de fonctions de celles-ci, à partir d'observations faites sur un premier échantillon et avec différents types de données.

Outre la densité prédictive, un prédicteur a été trouvé dans chaque cas et pour une distribution exponentielle. La loi a priori sur le paramètre est supposée de type conjuguée naturelle. On procède également à des simulations de données de type attribut dans le but de calculer des prédicteurs et des intervalles de prédictions de la 1<sup>ère</sup> statistique d'ordre.

**Abstract :** In this work, we consider the Bayesian prediction of order statistics and functions thereof from observations made on a first sample and with different types of data.

Besides the predictive density, a predictor has been found in each case and for an exponential distribution. We adopt the conjugate form for the prior distribution of the parameter. In order to compute predictors and predictive intervals of the first order statistic, we present some simulations whose data are of attribut type.<sup>1</sup>

## 1 Introduction

Le problème de la prédiction de statistiques d'ordre et de fonctions de celles-ci 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) [5]; Bancroft et Dunsmore (1976) [2]; Prezzi,

\*Université de Annaba BP 12, 23000 Algérie. e-mail : assiachadli428@hotmail.com

<sup>1</sup>Décision Bayésienne. *Mathematics Subject Classification.* 62M20, 62N05

Mots-clés : Statistique d'ordre, fiabilité, densité a priori, densité prédictive, prédicteur.

Geyskens et Monteiro (1996) [12]; Pan (1997) [11]; Tanner et Wong (1983) [14]; ainsi que Muller et Wang (1994) [9]. Dans le cas où les données sont censurées, le problème a été étudié par Padgett et McNichols (1984) [10]. L'approche Bayésienne a été également utilisée avec des lois a priori conjuguées naturelles ou non informatives sur des données complètes ou censurées et avec différentes lois. Bratcher-Schucany et Hunt (1971) [3] ont utilisé des données complètes; Dunsmore (1974) [4] a étudié d'une manière générale les tests de survie; Evans et Nigm (1980) [7, 8] ont étudié la loi exponentielle tronquée et la loi exponentielle bivariée avec des données censurées de type II; Arnold et Press (1989) [1] ont considéré la loi de Pareto; Erkanli, Mazuchi et Soyer (1998) [6] ainsi que Van-Batenberg et Kriens (1989) [15] ont appliqué les résultats de la prédiction Bayésienne à la fiabilité en particulier pour estimer le taux de panne et la fonction de fiabilité.

On dispose d'une expérience  $E$  dans laquelle  $n$  items sont soumis à un test de survie. La durée de vie de chaque item est  $X_i$  ( $i = 1, \dots, n$ ). Les  $X_i$  suivent une loi exponentielle à un paramètre dont la fonction de densité est

$$f(x | \theta) = \theta \exp(-\theta x), \quad x \geq 0, \theta > 0. \quad (1)$$

La fonction de répartition est

$$F(x | \theta) = 1 - \exp(-\theta x), \quad x \geq 0, \theta > 0.$$

On s'intéresse d'une part à la prédiction dans un futur échantillon  $(Y_1, \dots, Y_N)$ , indépendant du premier, de la  $k^{\text{ième}}$  statistique d'ordre

$$Y = Y_{(k)}, \quad 1 \leq k \leq N, \quad (2)$$

et de la durée de vie de  $N$  items jusqu'à la  $k^{\text{ième}}$  panne

$$Y = \sum_{i=1}^k Y_{(i)} + (N - k) Y_{(k)}, \quad (3)$$

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

$$Y = X_{(k)}, \quad r < k \leq n, \quad (4)$$

où  $0 \leq r \leq n$  est le nombre de pannes observé dans  $[0, T]$ ,

et

$$Y = \text{nombre de pannes dans } [X_{(r)}, X_{(r)} + T], \quad (5)$$

où  $(Y_{(1)}, Y_{(2)}, \dots, Y_{(N)})$  est la statistique d'ordre de  $(Y_1, Y_2, \dots, Y_N)$ ;  $X_{(r)}$  la  $r^{\text{ième}}$  statistique d'ordre;  $T > 0$  est un paramètre temps fixé d'avance.

Les prédictions Bayésiennes de (2), (3) pour des données censurées de type II ont été obtenus par Dunsmore (1974) [4] et Evans et Nigm (1980) [7], [8]. On se propose le même problème pour des données de type attribut.

Ce papier est organisé comme suit. En section 2, on décrit le problème avec quelques rappels. la section 3 donne la densité prédictive et un prédicteur de (2) et (3). En section 4, on s'intéresse à la prédiction dans un même échantillon de (4) et (5) quant on fixe le temps d'observation  $T$ . En section 5, on applique les résultats obtenus précédemment à la prédiction de la durée de vie de quelques systèmes. On termine en section 6 en procédant à des simulations pour la prédiction de  $Y_{(1)}$ .

## 2 Description du problème et préliminaires

On observe  $n$  variables aléatoires (v.a.)  $X_1, X_2, \dots, X_n$  indépendantes et identiquement distribuées (i.i.d.) de loi  $P^\theta$  et à valeurs dans  $(\mathcal{X}, \mathcal{B})$ , dans ce qui suit, on suppose que la loi  $P^\theta$  est la loi exponentielle de paramètre  $\theta$ . On suppose alors que les observations sont de type attribut pour prévoir (2), (3), (4) et (5).

### 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  $[0, T]$ ,  $x = 0, 1, \dots, n$ . La vraisemblance est alors

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

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

### 2.2 Distribution de la $k^{\text{ième}}$ statistique d'ordre

Soit  $(Y_1, Y_2, \dots, Y_N)$  un  $N$ -échantillon de même loi que  $X_1$ . La v.a.  $Y = Y_{(k)}$ ,  $1 \leq k \leq N$ , représente la durée de vie du  $k^{\text{ième}}$  item tombé en panne. On aurait besoin par la suite de la densité et de la fonction de répartition de

$Y = Y_{(k)}$  définis respectivement par :

$$f_{(k)}(y | \theta) = k \binom{N}{k} [F(y | \theta)]^{k-1} f(y | \theta) [1 - F(y | \theta)]^{N-k}, \quad y \geq 0,$$

et

$$F_{(k)}(y | \theta) = \sum_{i=k}^N \binom{N}{i} [F(y | \theta)]^i [1 - F(y | \theta)]^{N-i}.$$

### 2.3 Résolution de ce problème de décision par la méthode de Bayes

Le modèle statistique associé à ce problème de décision est  $(\mathcal{X}, \mathcal{B}, P_\theta)_{\theta \in \Theta}$ , où  $(P_\theta)_{\theta \in \Theta}$  est une famille de mesures de probabilité sur  $(\mathcal{X}, \mathcal{B})$  et  $\Theta$  est l'espace des paramètres. Soit  $Y$  le résultat d'une expérience future ayant pour densité de probabilité  $f(y/\theta)$ . Soit  $(D, \mathcal{D})$  l'ensemble des décisions muni de sa tribu Borélienne. La fonction de perte est supposée quadratique (voir Robert (1992) [13]). On pose

$$v : \Theta \times D \rightarrow \mathbb{R}^+, \quad v(\theta, d) = \int_{\mathbb{R}^+} (y - d)^2 f(y | \theta) dy.$$

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

On munit l'espace des paramètres  $\Theta$  d'une mesure de probabilité  $\tau$  de densité  $\pi(\theta)$  par rapport à la mesure de Lebesgue; la densité a posteriori de  $\theta$  est alors

$$\pi(\theta | x) = \frac{L(\theta; x) \pi(\theta)}{\int_{\Theta} L(\theta; x) \pi(\theta) d\theta},$$

où  $L(\theta; x)$  est la fonction de vraisemblance quant on a observé  $x$ . On calcule ensuite pour chaque décision  $d \in D$  le risque

$$R_\tau^x(d) = \int_{\Theta} v(\theta, d) \pi(\theta | x) d\theta = \int_{\mathbb{R}^+} (y - d)^2 p(y | x) dy = E((Y - d)^2 | x),$$

où

$$p(y | x) = \int_{\Theta} f(y | \theta) \pi(\theta | x) d\theta$$

est une densité appelée densité prédictive de  $Y$  quand  $x$  est observé. On cherche

$$\min_{d \in D} R_{\tau}^x(d),$$

celui-ci est égal à

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

où  $V(Y | x)$  est la variance prédictive de  $Y$ .

On conclut que la règle de décision  $\delta_b$  de Bayes associée à tout  $x \in \mathcal{X}$ , un prédicteur  $\delta_b(x)$  tel que :

$$\delta_b(x) = E(Y | x)$$

### 3 Prédiction dans le cas où les données sont de type attribut

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

La loi du nombre de pannes dans  $[0, T]$  est une loi binomiale  $b(n, p)$  où la probabilité d'un succès est  $p = 1 - \exp(-\theta T)$ , d'où la vraisemblance

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

La loi a priori sur  $\theta$  étant une loi *gamma* ( $g, h$ ), la densité a posteriori est alors donnée par

$$\pi(\theta | x) = \frac{1}{A\Gamma(g)} \sum_{i=0}^x (-1)^i \binom{x}{i} \theta^{g-1} \exp(-\theta(H+iT)), \quad \theta > 0,$$

où  $\Gamma(g) = \int_0^\infty x^{g-1} e^{-x} dx$ ,  $H = h + T(n-x)$ ,  $A = \sum_{i=0}^x (-1)^i \binom{x}{i} (H+iT)^{-g}$ .

#### 3.1 Prédiction de $Y = Y_{(k)}$ , $1 \leq k \leq N$

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

$$f_{(k)}(y | \theta) = k \binom{N}{k} \sum_{j=0}^{k-1} (-1)^j \binom{k-1}{j} \theta \exp(-\theta y(N-k+1+j)), \quad y \geq 0.$$

L'intégration par rapport à  $\theta$  du produit  $f_{(k)}(y | \theta) \pi(\theta | x)$  conduit à la densité prédictive

$$\begin{aligned} p(y | x) &= \frac{k \binom{N}{k}}{A \Gamma(g)} \sum_{i=0}^x \sum_{j=0}^{k-1} (-1)^{i+j} \binom{x}{i} \binom{k-1}{j} \\ &\quad \times \int_0^\infty \theta^g \exp(-\theta(y(N-k+1+j) + H + iT)) d\theta \\ &= \frac{gk \binom{N}{k}}{A} \sum_{i=0}^x \sum_{j=0}^{k-1} (-1)^{i+j} \binom{x}{i} \binom{k-1}{j} \\ &\quad \times [y(N-k+1+j) + H + iT]^{-(g+1)} \end{aligned} \quad (6)$$

La fonction de répartition prédictive de  $Y$  est

$$\begin{aligned} F(y | x) &= \int_0^y p(z | x) dz \\ &= \frac{gk \binom{N}{k}}{A} \sum_{i=0}^x \sum_{j=0}^{k-1} (-1)^{i+j} \binom{x}{i} \binom{k-1}{j} \\ &\quad \times \int_0^y [z(N-k+1+j) + H + iT]^{-(g+1)} dz \\ &= \frac{k \binom{N}{k}}{A} \sum_{i=0}^x \sum_{j=0}^{k-1} (-1)^{i+j} \binom{x}{i} \binom{k-1}{j} (N-k+1+j)^{-1} \\ &\quad \times \{ (H + iT)^{-g} - [y(N-k+1+j) + H + iT]^{-g} \}. \end{aligned}$$

En particulier pour  $k = 1$

$$F(y | x) = 1 - \frac{1}{A} \sum_{i=0}^x (-1)^i \binom{x}{i} [Ny + H + iT]^{-g}.$$

La fonction  $p(y | x)$  étant une fonction strictement 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

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

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

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

$$\begin{aligned}
 E(Y | x) &= \int_0^\infty yp(y | x) dy \\
 &= \frac{k \binom{N}{k}}{A(g-1)} \sum_{i=0}^x \sum_{j=0}^{k-1} (-1)^{i+j} \binom{x}{i} \binom{k-1}{j} \frac{(N-k+1+j)^{-2}}{(H+iT)^{g-1}} \\
 &= \frac{k \binom{N}{k}}{A(g-1)} \left[ \sum_{i=0}^x (-1)^i \binom{x}{i} (H+iT)^{1-g} \right] \\
 &\quad \times \left[ \sum_{j=0}^{k-1} (-1)^j \binom{k-1}{j} (N-k+1+j)^{-2} \right].
 \end{aligned} \tag{7}$$

En particulier pour  $k = 1$  :

$$E(Y_{(1)} | x) = \frac{1}{A(g-1)N} \sum_{i=0}^x (-1)^i \binom{x}{i} (H+iT)^{1-g}. \tag{8}$$

### 3.2 Prédiction de $Y = \sum_{i=1}^k Y_{(i)} + (N-k) Y_{(k)}$

Du fait de la propriété de perte de mémoire de la loi exponentielle, il est connu que si  $Y_{(1)}, Y_{(2)}, \dots, Y_{(N)}$  est la statistique d'ordre de  $Y_1, Y_2, \dots, Y_N$  où les v.a.  $Y_i$  ( $i = 1, \dots, N$ ) sont i.i.d. et de loi  $\exp(\theta)$  alors  $Y$  suit une loi  $\text{gamma}(k, \theta)$  (voir Dunsore [4]) dont la densité est :

$$f(y | \theta) = \frac{\theta^k}{\Gamma(k)} y^{k-1} \exp(-\theta y), \quad y \geq 0, \theta > 0.$$

La densité prédictive  $p(y | x)$  de  $Y$  s'écrit :

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

où

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

On remarque que la loi conditionnelle de  $Y$  sachant  $x$  est une combinaison linéaire de lois beta de type II, où la loi  $\beta_{II}(k, g)$  est définie par sa densité :

$$f(y) = \frac{1}{\beta(k, g)} \times \frac{y^{k-1}}{(1+y)^{k+g}}, \quad y \geq 0$$

La fonction de répartition prédictive  $F(y | x)$  de  $Y$  est

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

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

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

$$I_{(a_i)}(k, g) = \frac{1}{\beta(k, g)} \int_0^{a_i} \frac{y^{k-1}}{(1+y)^{k+g}} dy,$$

où  $I_{(a_i)}(k, g)$  est la valeur d'une loi beta de type II de paramètres  $k$  et  $g$  au point  $a_i = z / (H + iT)$ . Sous l'hypothèse  $g > 1$ , un prédicteur de  $Y$  serait

$$\begin{aligned} E(Y | x) &= \frac{1}{A\beta(k, g)} \sum_{i=0}^x (-1)^i \binom{x}{i} \int_0^\infty \frac{y^k}{[y + H + iT]^{k+g}} dy \quad (9) \\ &= \frac{\beta(k+1, g-1)}{A\beta(k, g)} \sum_{i=0}^x (-1)^i \binom{x}{i} (H + iT)^{1-g} \\ &= \frac{k}{A(g-1)} \sum_{i=0}^x (-1)^i \binom{x}{i} (H + iT)^{1-g}. \end{aligned}$$

On remarque que pour  $k = 1$ ,  $E(Y | x) = NE(Y_{(1)} | x)$ , (voir (8)).

## 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  $Y = X_{(r+l)} - T$ , ( $l = 1, 2, \dots, n - r$ ) et à celle du nombre de pannes  $Y$  dans  $[T, T + t]$ , où  $t > 0$  est un temps fixé à l'avance.

#### 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  $x > T$  :

$$f_{(r+l)}(x | \theta) = \frac{1}{\beta(l, n-r-l+1)} \left[ \frac{F(x | \theta) - F(T | \theta)}{1 - F(T | \theta)} \right]^{l-1} \frac{f(x | \theta)}{1 - F(T | \theta)} \\ \times \left[ \frac{1 - F(x | \theta)}{1 - F(T | \theta)} \right]^{n-r-l}.$$

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

$$f_{(l)}(y | \theta) = l \binom{n-r}{l} \sum_{j=0}^{l-1} (-1)^j \binom{l-1}{j} \theta \exp(-\theta y (n-r-l+j+1)), \quad y \geq 0.$$

En remarquant que  $Y = X_{(r+l)} - T$  a la même distribution que la  $l^{\text{ème}}$  statistique d'ordre d'un échantillon de taille  $n-r$ , les résultats de cette sous-section se déduisent de ceux de la sous-section 3.1 en remplaçant  $k$  par  $l$ ,  $N$  par  $n-r$  et  $x$  par  $r$ . La densité a priori sur  $\theta$  est supposée celle d'une loi *gamma*  $(g, h)$ , la densité prédictive  $p(y | r)$  est alors

$$p(y | r) = \frac{g l \binom{n-r}{l}}{A} \sum_{i=0}^r \sum_{j=0}^{l-1} (-1)^{i+j} \binom{r}{i} \binom{l-1}{j} \\ \times [y(n-r-l+j+1) + H + iT]^{-(g+1)}.$$

En particulier pour  $l = 1$

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

La fonction de répartition prédictive  $F(y | r)$  de  $Y$  est alors

$$F(y | r) = \frac{l \binom{n-r}{l}}{A} \sum_{i=0}^r \sum_{j=0}^{l-1} (-1)^{i+j} \binom{r}{i} \binom{l-1}{j} (n-r-l+j+1)^{-1} \\ \times \{ (H + iT)^{-g} - [y(n-r-l+j+1) + H + iT]^{-g} \}.$$

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

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

Par rapport à une fonction de perte quadratique et sous l'hypothèse  $g > 1$  (voir (7))

$$E(Y | r) = \frac{l \binom{n-r}{l}}{A(g-1)} \left[ \sum_{i=0}^r (-1)^i \binom{r}{i} [H + iT]^{-(g-1)} \right] \\ \times \left[ \sum_{j=0}^{l-1} (-1)^j \binom{l-1}{j} (n-r-l+1+j)^{-2} \right].$$

En particulier pour  $l = 1$  (voir (8))

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

où  $Y_1 = X_{(r+1)} - T$ .

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

Soit  $Y$  la v.a. donnant le nombre de pannes dans  $[T, T+t]$ . Puisque

$$(Y > l) \Leftrightarrow [X_{(r+l+1)} - T \leq t],$$

on a

$$P[(Y > l) | r] = P(X_{(r+l+1)} - T \leq t | r).$$

On connaît la loi de  $X_{(r+l)}$  sachant  $r$ , on peut en déduire la loi de  $X_{(r+l+1)}$  sachant  $r$  par

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

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  $E(Y | r)$  :

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

## 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 (1). A partir d'observations censurées de type II ou d'observations de type attribut d'un  $n$ -échantillon de (1), les résultats obtenus en sections 3.1, 3.2 et 4.1 permettent la prédiction de la durée de vie du système  $S_k$ . Ci dessous quelques exemples de systèmes.

### 5.1 “ $r$ out of $k$ system”

Etant donné un système  $S_k$  qui fonctionne si et seulement si  $r$  des  $k$  composants fonctionnent simultanément.

“ $r$  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  $Y_1, Y_2, \dots, Y_k$  de chacun des composants du système constituent un  $k$ -échantillon de (1) indépendant de l'échantillon  $(X_1, X_2, \dots, X_n)$  des observations. Ainsi la prédiction de  $Y$  est équivalente à la prédiction de  $Y_{(r)}$ ,  $1 \leq r \leq k$ . On peut donc appliquer les résultats de la sous-section 3.1 avec  $N = k$  et  $k = r$  dans le cas de données de type attribut.

### 5.2 “Stand-by redundant system”

Etant donné un système  $S_k$  à  $k$  composants identiques; un composant d'origine et  $(k-1)$  composants de rechange, chacun commençant son fonc-

tionnement 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 :

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

On peut donc appliquer les résultats de la sous-section section 3.2 avec  $k = N$ .

### 5.3 Système en série avec "stand-by components"

Un tel système est constitué de  $k$  composants opérant en série avec  $(r - 1)$  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 des sous-sections 3.1 et 4.1.

## 6 Simulations

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  $\theta$ , on calcule la fonction de risque  $R(\theta, d(x))$  en utilisant la loi des grands nombres. Un prédicteur de  $Y = Y_{(1)}$  quand  $x$  est observé est obtenu en (8).

Un prédicteur de  $Y = \sum_{i=1}^k Y_{(i)} + (N - k) Y_{(k)}$  quand  $x$  est observé est obtenu en (9).

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 :  $n = 10$  le nombre d'items observés dans  $[0, T]$ ,  $T = 10$  la durée de l'observation,  $g = 2$  et  $h = 0.01$  sont les paramètres de la loi a priori sur  $\theta$  qui est de type  $\text{gamma}(g, h)$ ,  $N = 20$  taille d'un échantillon futur tiré d'une même loi  $\exp(\theta)$ ;  $x$  le nombre de défaillances (noté  $ix$ ) observés dans  $[0, T]$ ;  $x = 0, 1, \dots, n$ . Les valeurs de  $x$  sont simulées.

Les résultats pour chaque valeur de  $\theta$  sont groupés par tableau : 1ère ligne : les valeurs de ix (simulées), 2ème ligne : les valeurs de la première statistique d'ordre  $Y_{(1)}$  d'un échantillon futur indépendant du premier et de taille  $N$  (simulées), 3ème ligne : les valeurs pour chaque ix observé d'un prédicteur  $d(x)$  de  $Y_{(1)}$  (calculées en utilisant un programme fortran), 4ème ligne : on calcule les écarts quadratiques  $(Y_{(1)} - d(x))^2$ , 5ème ligne : pour chaque ix observé, on calcule une borne supérieure de prédiction  $z$  de niveau 0.95 et avec une précision  $e = 0.001$  ( $z$  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  $Y_{(1)}$  ( $MT = 1/N\theta$ ), la moyenne des prédicteurs MP et la valeur de la fonction de risque au point  $\theta$  :

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

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

Table 1 : Résultats de la simulation pour  $\theta = 0.0142$

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

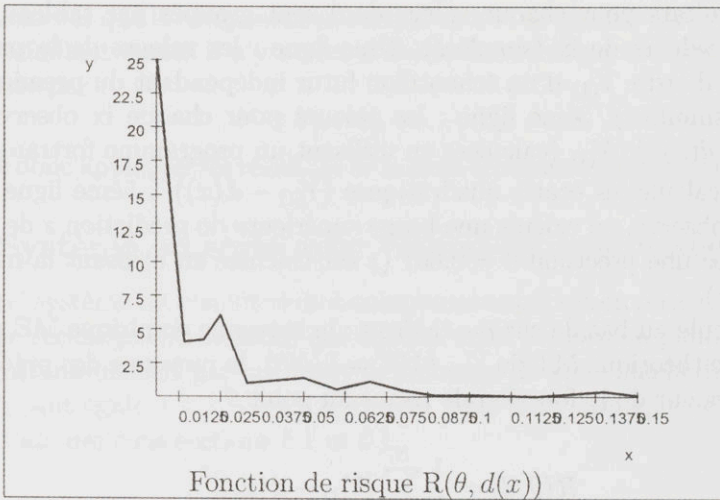
Valeurs calculées à partir de la table 1 :

ME=2.75; MT=3.52; MP=2.64;  $R(\theta, d(x)) = 4.05$ .

On refait ce travail pour différentes valeurs de  $\theta$  afin de tracer la fonction de risque.

Valeurs de  $\theta$  choisies pour tracer  $R(\theta, d(x))$  :

0.0060	0.0110	0.0125	0.0142	0.0200	0.0250	0.0330	0.0500
0.0600	0.0700	0.0800	0.0900	0.1000	0.1200	0.1400	0.1500



## 7 Conclusion

On remarque qu'en moyenne pour des valeurs de  $g$  et  $h$  se rapprochant de ceux d'une loi a priori non informative sur  $\theta$ , la valeur d'un prédicteur de  $Y_{(1)}$  est sous estimé par rapport à sa valeur moyenne théorique.

Ceci est particulièrement visible pour les valeurs de  $\theta$  telles que la moyenne de durée de vie ( $1/\theta$ ) soit grande par rapport au temps d'observations  $T$ .

La fonction de risque  $R(\theta, d(X))$  présente de grandes variations pour les petites valeurs de  $\theta$  (i.e. pour des moyennes de durées de vie élevées par rapport au temps d'observation  $T$ ); ceci est a priori du à un nombre de simulations trop petit

## Références

- [1] Arnold B.C. et Press S.J. *Bayesian estimation and prediction for Pareto data*. J.A.S.A., Vol 84, N°408 (1989) 1079-1084.
- [2] Bancroft G.A. et Dunsmore I.R. *Prédictive distribution in life tests under competing causes of failure*. Biometrika, Vol 63, N°1 (1976) 195-217
- [3] Bratcher T., Schucany W. et Hunt H. *Bayesian prédiction and population size assumptions*. Technometrics, Vol 13, N°3 (1971) 678-681.

- [4] Dunsmore I.R. *The Bayesian predictive distribution in life testing models*. Technometrics, Vol 16, N°3 (1974) 455-460.
- [5] Dunsmore I.R. *Asymptotic prediction analysis*. Biometrika, Vol 63, (1976) 627-630.
- [6] Erkanli A., Mazzuchi T. et Soyer R. *Bayesian computation for a class of reliability growth models*. Technometrics, Vol 40, N°1 (1998) 14-23.
- [7] Evans I.G. et Nigm A.H.M. *Bayesian prediction for the left truncated exponential distribution*. Technometrics, Vol 22, N°2 (1980) 329-340.
- [8] Evans I.G. et Nigm A.H.M.. *Bayesian 1-sample prediction for the 2-parameter distribution*. I.E.E.E. transactions on reliability, Vol R-29, N°5 (1980) 218-229.
- [9] Muller H.G. et Wang J.L.. *Hazard rate estimation under random censoring with varying kernels and bandwidths*. Biometrics, Vol 50, N°3 (1994) 61-76.
- [10] Padgett W.J. et McNichols D.T. *non parametric estimation from censored data*. Comm.statist.theory methods Vol 13 (1984) 1581-1613.
- [11] Pan J.N. *Reliability prediction of imperfect survitching systems subject to multiple stress*. Microelectronics and reliability. ISSN0026-2714, Vol 37, N°3 (1997) 439-445.
- [12] Prezzi M., Geyskens P. et Monteiro P.M. *Reliability approach to survive life prediction of concrete exposed to marine environments*. ACI materials journal. ISSN 0889-325X, Vol 93, N°6 (1996) 544-552.
- [13] Robert C. *l'analyse statistique Bayesienne*. (1992) Economica.
- [14] Tanner M. et Wong W. *The estimation of the hazard function from randomly censored data by the kernel method*. Ann. statist. Vol 11, N°3 (1983) 989-993.
- [15] Van-Batenberg P.C. et Kriens J. *Bayesian discovery sampling a simple model of Bayesian inference in auditing*. Statistician. ISSN 0039-0526, Vol 38, N°4 (1989) 227-233.