



## Quelques considérations sur l'utilisation pratique des modèles discrets de survie en fiabilité industrielle

Alberto Pasanisi, Côme Roero, Gilles Celeux, Emmanuel Rémy

### ► To cite this version:

Alberto Pasanisi, Côme Roero, Gilles Celeux, Emmanuel Rémy. Quelques considérations sur l'utilisation pratique des modèles discrets de survie en fiabilité industrielle. 41èmes Journées de Statistique, SFdS, Bordeaux, 2009, Bordeaux, France, France. inria-00386574

**HAL Id: inria-00386574**

**<https://hal.inria.fr/inria-00386574>**

Submitted on 22 May 2009

**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.

# QUELQUES CONSIDÉRATIONS SUR L'UTILISATION PRATIQUE DES MODÈLES DISCRETS DE SURVIE EN FIABILITÉ INDUSTRIELLE

Alberto Pasanisi\*; Côme Roero\*\*, Gilles Celeux\*\*, Emmanuel Rémy\*

\* EDF-R&D, Management des Risques Industriels, 6 quai Watier, 78401 Chatou Cedex.  
alberto.pasanisi@edf.fr, emmanuel.remy@edf.fr

\*\* INRIA Futurs, Université Paris-Sud, Dept. de mathématiques, Bât. 425, 91405 Orsay Cedex  
gilles.celeux@math.u-psud.fr, croero@hotmail.com

## Résumé

Les modèles discrets de survie sont a priori à utiliser dans le cas de données discrètes de durée de vie. Des données de ce type peuvent provenir, par exemple, d'un matériel sollicité seulement par moments. La variable à modéliser est donc celle correspondant au numéro de la sollicitation à laquelle il y a défaillance.

Nous nous sommes intéressés dans le cadre de cette étude aux deux modèles "Polya inverse", basé sur le schéma d'urne de Polya, et "Weibull 1", un des possibles analogues discrets du modèle continu de Weibull.

Après en avoir présenté les principales propriétés mathématiques, nous mettons en évidence une faiblesse importante du modèle "Polya inverse" : il ne peut modéliser correctement que des matériels dont le vieillissement est décéléré. Cette contrainte limite considérablement son intérêt pratique.

Le modèle "Weibull 1", plus flexible, ne souffre pas de ce problème. Néanmoins, il présente une grande ressemblance avec le modèle continu de Weibull, ressemblance qui est de plus en plus forte en présence de données de survie de valeur élevée (c'est-à-dire relatives à des matériaux qui défont au bout d'un grand nombre de sollicitations) ou censurées. Dans ces situations, qui sont celles qui se rencontrent le plus fréquemment dans un contexte industriel comme celui d'EDF, le modèle "Weibull 1" est bien approximé, et donc remplacé, par son analogue continu, mieux appréhendé par les ingénieurs et dont le maniement est beaucoup plus aisé.

Mots-clé : Données discrètes de survie, Modèle de Polya inverse, Modèle de Weibull discret, vieillissement

## Abstract

Discrete lifetime models are a priori to be used in the case of discrete survival data. Such kind of data can come, for example, from the survey of a component which operates on demands only. The variable to be modelled is thus related to the number of demands prior to failure.

We were interested within the framework of this study in the two models "Polya inverse", based on the Polya's urn scheme, and "Weibull 1", one of the possible discrete variants of the continuous Weibull model.

After having presented the principal mathematical properties of them, we highlight an important weakness of the "Polya inverse" model: it can model correctly only components the ageing of which is decelerated. This constraint weakens considerably its practical interest.

The "Weibull 1" model, more flexible, is not affected by this problem. Nevertheless, it presents a large resemblance to the continuous Weibull model, resemblance which strongly increases in the presence of high value survival data (i.e. relative to materials which fail at the end of a great number of requests) or censored data. In these situations, which are the ones most frequently met in an industrial context as EDF's one, the "Weibull 1" model can be approximated, and replaced, by

continuous Weibull model, better understood by engineers and the handling of which is much easier.

Key-words : Discrete survival data, Inverse Polya model, Discrete Weibull model, ageing.

## Introduction

L'utilisation des modèles discrets de survie est envisageable dans les cas où la durée de vie d'un matériel ne peut pas être décrite par le temps calendaire. Ces situations se rencontrent fréquemment dans la pratique industrielle.

C'est par exemple le cas de composants qui ne fonctionnent pas en continu mais dont le fonctionnement est appelé ponctuellement (par exemple disjoncteurs d'arrêt automatique, groupes électrogènes). L'appel au composant, par exemple le démarrage d'un groupe électrogène, peut être interprété comme une "sollicitation". Pour cette raison, ces modèles sont aussi couramment appelés "modèles de défaillance à la sollicitation". La variable aléatoire que nous cherchons ici à décrire est alors celle correspondant au numéro  $N$  de la sollicitation à laquelle il y a défaillance du matériel. En d'autres termes,  $N$  est le nombre total d'utilisations du matériel avant sa défaillance.

Un autre cas pratique de données discrètes de survie se présente quand le bon fonctionnement d'un matériel en fonctionnement continu n'est relevé que par périodes. Les données à modéliser sont, dans ce cas, le nombre de périodes pendant lesquelles le composant a bien fonctionné avant sa défaillance.

Les modèles discrets de survie ont été relativement peu étudiés, par rapport aux modèles continus, dont l'utilisation est courante dans la pratique industrielle. A titre de comparaison, le premier article qui propose un modèle spécifique (dédit du modèle continu de Weibull) pour des données discrètes de survie date seulement de 1975 (Nagakawa et Osaki S, 1975). Un résumé exhaustif des différents modèles discrets est fourni par Bracquemont et Gaudoin (2003) qui regroupent ces modèles en deux grandes familles : ceux qui sont directement dérivés de modèles continus de durée de vie et ceux dérivés du modèle d'urne de Polya.

Nous présentons ici les résultats d'un travail mené conjointement par EDF-R&D et l'INRIA sur ce sujet. Le but de ce travail était d'acquérir une meilleure compréhension des modèles discrets de survie et d'analyser leur pertinence dans le contexte appliqué de la fiabilité industrielle.

Ces travaux s'inscrivent dans le cadre d'une collaboration pluriannuelle entre ces deux partenaires dans le domaine des modèles de survie (Bousquet, 2005), (Bousquet et al., 2006).

## Les modèles étudiés

Dans le cadre de cette étude, nous nous sommes focalisés sur les deux modèles "Polya Inverse" et "Weibull 1".

Dans le modèle "Polya Inverse", chaque sollicitation du composant est assimilée à un tirage dans une urne contenant initialement  $a$  boules blanches (associées à l'évènement "bon fonctionnement") et  $b$  boules noires (associées à l'évènement "défaillance"). Après chaque tirage, si la boule sortie est blanche, alors on rajoute  $z$  boules noires, augmentant ainsi la probabilité de défaillance lors du tirage suivant. Si la boule sortie est noire, on arrête les tirages : en effet, nous nous intéressons ici à la première défaillance.

Le modèle "Polya Inverse" est paramétré par les deux grandeurs :

- $\beta$  : rapport entre le nombre de boules noires ( $b$ ) et le nombre total de boules dans l'urne ( $a+b$ ) lors du premier tirage ;
- $\zeta$  : rapport entre le nombre de boules noires rajoutées ( $z$ ) après chaque tirage d'une boule blanche et le nombre initial de boules ( $a+b$ ).

Le modèle « Weibull 1 » est un des analogues discrets du modèle classique de Weibull « continu ».

Il est défini par sa fonction de survie :

$$R(n) = e^{(-n/\eta)^\beta} \text{ où } n \text{ est entier,}$$

parfois écrite sous la forme :

$$R(n) = q^{n^\beta} \text{ avec } q = e^{(-1/\eta^\beta)}.$$

Il s'agit de deux modèles à deux paramètres :  $(\beta, \zeta)$  pour "Polya Inverse" et  $(\eta, \beta)$ , ou  $(q, \beta)$ , pour « Weibull 1 ».

Bien que d'autres définitions sont possibles (Xie et al., 2005), le taux de défaillance des modèles discrets est défini usuellement par la probabilité conditionnelle :

$$p(n) = P(N = n | N > n - 1),$$

ce qui donne pour les deux modèles en question les deux formulations suivantes :

$$p(n) = \frac{\beta + (n-1)\zeta}{1 + (n-1)\zeta} \quad \text{Polya Inverse}$$

$$p(n) = 1 - \exp\left(-\left(\frac{n}{\eta}\right)^\beta + \left(\frac{n-1}{\eta}\right)^\beta\right) \quad \text{Weibull 1}$$

## Principaux résultats et enseignements

Dans le modèle "Polya Inverse", puisque le nombre de boules noires rajoutées dans l'urne après chaque tirage d'une boule blanche (non-défaillance) est constant, le taux de défaillance est une fonction, certes, croissante du nombre de sollicitations  $N$ , mais marginalement décroissante.

En effet, plus le nombre de sollicitations est élevé, plus le nombre de boules noires déjà présentes dans l'urne est élevé, et donc la contribution apportée au taux de défaillance par les  $z$  boules noires rajoutées sera de plus en plus faible. Cette intuition a été démontrée mathématiquement.

Le taux de défaillance de "Polya Inverse" est une fonction forcément concave qui s'approche d'une fonction linéaire pour des petites valeurs du paramètre  $\zeta$  (cf. figure 1).

Des exercices d'estimation ont montré que, en présence de données issues de situations de vieillissement accéléré, le modèle "Polya Inverse", ne pouvant pas décrire ces phénomènes, essaie de s'adapter au mieux aux données, en faisant tendre son taux de défaillance vers son comportement "limite" linéaire (ce qui correspond à des valeurs très petites de  $\zeta$ ).

Pour cette raison, malgré son caractère intuitif et séduisant (Clarotti et al., 1997), ce modèle ne nous semble pas adapté au contexte industriel, où il est bien difficile d'imaginer des matériels dont le vieillissement est par définition décéléré.

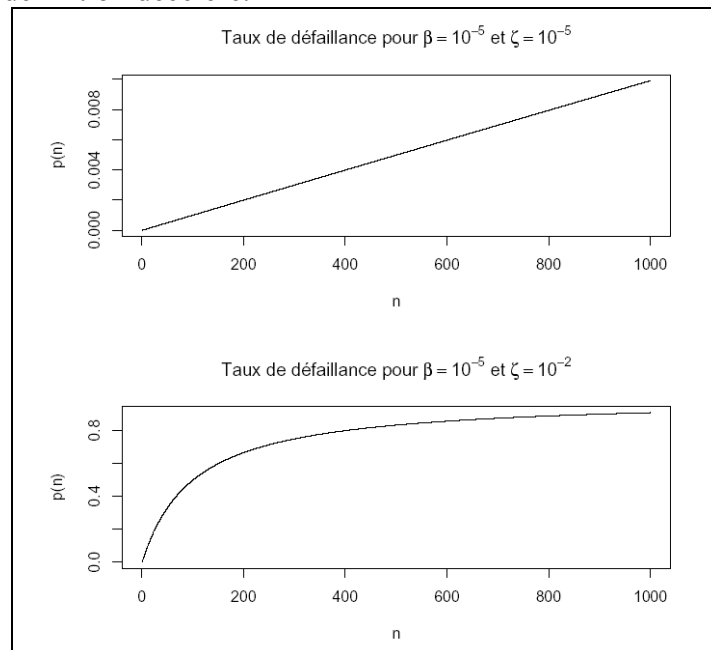


Fig. 1. Taux de défaillance du modèle « Polya Inverse » pour deux couples de valeurs des paramètres  $\beta$  et  $\zeta$

À propos du modèle "Weibull 1", nous avons tout d'abord étudié les propriétés du taux de défaillance pour voir dans quelle mesure les relations entre la valeur du paramètre de forme  $\beta$  et le type de vieillissement (bien connues et appréhendées dans la pratique technique) pouvaient être étendues au cas discret.

Une différence est que dans le cas discret, le taux de défaillance est une probabilité, alors que dans le cas continu, il s'agit du rapport entre une densité de probabilité et une probabilité et donc, d'une fonction non bornée. Par conséquent, le taux de défaillance dans le cas discret est soit une fonction concave de  $n$ , soit une fonction en forme de S. Les situations de vieillissement accéléré (ou plutôt non décéléré) sont décrites par un taux de défaillance initialement convexe, puis concave et caractérisé par la présence d'un point d'inflexion.

Pour  $\beta \leq 2$ , il est possible de démontrer analytiquement que le taux de défaillance est une fonction concave de  $N$  (représentative de situations de vieillissement décéléré).

Pour  $\beta > 2$ , la courbe des taux de défaillance est soit strictement concave pour de petites valeurs de  $\eta$ , soit elle présente un point d'inflexion pour de grandes valeurs de  $\eta$ .

En fonction de  $\eta$  et de  $\beta$ , nous avons évalué la valeur de  $N$  pour laquelle la fonction  $p''$

$$p''(n) = p(n) - 2p(n-1) + p(n-2),$$

pouvant être interprétée comme un équivalent discret de la dérivée seconde du taux de défaillance, change de signe (point d'inflexion). Une conclusion importante se dégage : dans le cas du modèle discret "Weibull 1", le type de vieillissement (accéléré ou décéléré) ne dépend pas uniquement de  $\beta$ , comme dans le cas continu, mais aussi de  $\eta$ .

Les modèles "Weibull continu" et "Weibull 1" étant très proches, nous nous sommes posés ensuite la question de l'intérêt pratique de l'utilisation du modèle discret, par rapport au modèle continu (plus facile à estimer et à appréhender) en présence de données de défaillance à la sollicitation. En particulier, nous avons simulé plusieurs jeux de données à partir du modèle "Weibull 1", en utilisant des valeurs différentes des paramètres  $\eta$  et  $\beta$  et différents taux de censure, et nous avons estimé par maximum de vraisemblance ces paramètres, selon les deux hypothèses de modèle sous-jacent : "Weibull 1" et "Weibull continu".

Ensuite, nous avons analysé et comparé, suivant les deux hypothèses ci-dessus, les écarts entre les estimations de  $(\eta, \beta)$  et les valeurs réelles, ainsi que les mêmes écarts relatifs à d'autres grandeurs d'intérêt fiabiliste (nombre moyen de sollicitations avant défaillance, taux de défaillance, quantiles de la variable discrète "durée de vie").

Dans la plupart des cas, les valeurs estimées à partir de l'hypothèse de modèle continu sont très proches des valeurs réelles (issues du modèle discret), surtout pour des valeurs élevées de  $\eta$ , représentatives de matériels qui tombent en panne après un nombre élevé de sollicitations.

Cela est de plus en plus vrai à mesure que le taux de censure augmente. En effet, une observation censurée se traduit par l'apport d'un terme à la vraisemblance qui est exactement le même dans le cas discret et dans le cas continu : plus il y a de données censurées, plus les deux vraisemblances "discrète" et "continue" seront similaires.

Compte tenu de ces résultats, il ressort que l'intérêt du modèle "Weibull 1" en fiabilité industrielle est modeste. Dans la plupart des cas pratiques pouvant intéresser EDF, le modèle discret peut être remplacé par son équivalent continu, bien appréhendé et de manipulation plus aisée.

Les modèles discrets de vieillissement ne sauraient être utiles que pour des données de défaillance de valeurs très faibles et très peu nombreuses. Outre que ces cas sont rarissimes, du moins dans le contexte industriel d'EDF, on peut légitimement se demander s'ils doivent donner lieu à une modélisation, alors que des simples valeurs empiriques pourraient nous renseigner sur la présence de vieillissement ou non.

## Bibliographie

[1] Nagakawa, T. et Osaki S. (1975) The discrete Weibull distribution. *IEEE Transactions on Reliability*, 34 (5), 300-301.

- [2] Bracquemont, C. Gaudoin, O. (2003) A survey on discrete lifetime distributions. *International Journal of Reliability, Quality and Safety Engineering*, 10 (1), 69-98.
- [3] Bousquet N. (2005) Eliciting prior distributions for Weibull inference in an industrial context. *Communications in Dependability and Quality Management*, 8 (3).
- [4] Bousquet, N. Celeux, G., Billy, F. Remy, E. (2006) Inférence des paramètres d'une loi de Weibull - Approches classique et bayésienne, *Lambda-Mu 15*, Lille, France.
- [5] Xie, M. Gaudoin, O. Bracquemont, C. (2002) Redefining failure rate functions for discrete distributions. *Journal of Reliability, Quality and Safety Engineering*, 9 (3), 275-285.
- [6] Clarotti, A, Lannoy, A. et Procaccia, H. (1997) Probability risk analysis of ageing components which fail on demand : a Bayesian model. Application to maintenance optimization of diesel engine linings. Dans *Ageing of Materials and Methods for the Assesment of Lifetimes of Engineering Plant*. Penny ed., Balkema, Rotterdam, pp 85-93.