HAL
open science

Critères de sélection de modèle de fiabilité dans le cadre des risques concurrents

Yann Dijoux

To cite this version:

Yann Dijoux. Critères de sélection de modèle de fiabilité dans le cadre des risques concurrents. 41èmes Journées de Statistique, Société Française de Statistique, May 2009, Bordeaux, France. inria-00386710

HAL Id: inria-00386710

https://inria.hal.science/inria-00386710v1

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.

Critères de sélection de modèle de fiabilité dans le cadre des risques concurrents

Yann Dijoux

Management Science Department, University of Strathclyde, Glasgow

Résumé

Afin de modéliser la dépendance entre les maintenances correctives et préventives conditionnelles, nous utilisons le cadre des risques concurrents. Dans ce contexte, la fonction ϕ peut être aisément estimée et est très utile lors de la sélection de modèle. Dans cette communication, nous proposons dans un premier temps d'étudier les propriétés de l'estimateur non paramétrique de ϕ. Enfin, nous présentons un test de monotonie de la fonction ϕ. Nous illustrons les résultats obtenus à l'aide d'un jeu de données réelles.

Abstract

The competing risks framework is used for the joint modelling of the dependency between corrective and condition-based preventive maintenance. In this context, the ϕ function can be easily estimated and is very useful in model selection. In this communication, we first study the quality of the non-parametric estimator of ϕ. Finally, we present a statistical test on the monotony of ϕ. The main results are illustrated by the study of a real dataset.

Mots clés : fiabilité, risque industriel, processus aléatoires ponctuels, risques concurrents.

1. Introduction

Tout au long de leur vie opérationnelle, les systèmes complexes (centrales nucléaires, automobiles, etc...) sont soumis à deux types de maintenance:

- Les maintenances correctives (MC) interviennent suite à une défaillance et visent à remettre le système en marche.
- Les maintenances préventives (MP) interviennent lorsque le système est en état de fonctionnement et ont pour but de réduire la dégradation du système et retarder une éventuelle panne ultérieure. Nous nous intéressons plus particulièrement aux MP conditionnelles, effectuées suite à l'observation d'une dégradation avancée du système par l'équipe de maintenance.

Pour un système neuf, on ne sait pas si la première maintenance sera corrective ou préventive. Soit Z le temps potentiel d'attente jusqu'à la première MC suite à une défaillance. Cette panne peut-être potentiellement évitée par une maintenance préventive

à l'instant Y. Ainsi, le temps réel d'attente avant la première maintenance est $W = \min(Y, Z)$ et le type de cette maintenance est déterminé par $U = \mathbb{1}_{\{Y \leq Z\}}$. $U = 0$ si la maintenance est corrective et $U = 1$ si elle est préventive.

Nous venons de présenter le principe des risques concurrents (Crowder, 2001), développé dans le contexte de la maintenance par Cooke et Paulsen (1994) ou Cooke et Bedford (2002). Désormais, Y et Z seront respectivement appelés risque de MP et risque de MC.

Par la suite, nous définissons la fonction ϕ, couramment employée dans le cadre des risques concurrents. Puis, nous présentons les principaux modèles de risques concurrents en précisant la monotonie de la fonction ϕ pour chacun des modèles. L' estimateur non paramétrique de ϕ sera alors présenté ainsi que de nouveaux résultats sur la qualité de cet estimateur. Enfin, nous présentons un test de monotonie sur ϕ basé sur la statistique de Wilcoxon.

2. Le contexte des risques concurrents

2.1 Définitions

La dépendance entre le risque de MP Y et le risque de MC Z s'exprime par la loi jointe de (Y, Z), définie par sa fonction de survie bivariée $S(y, z) = P(Y > y, Z > z)$. Les difficultés d'identifiabilité font que l'on ne peut estimer que les fonctions liées à la loi jointe de (W, U). En particulier, on définit une fonction appelée communément ϕ:

$$\phi(w) = P(Y \leq Z | W > w) = P(U = 1 | W > w)$$

Concrètement, $\phi(w)$ est la probabilité que la prochaine maintenance soit préventive sachant qu'elle n'a pas encore eu lieu à l'instant w. Etant estimable, cette fonction permet d'avoir des critères sur le choix d'un modèle de risques concurrents approprié à un jeu de données particulier.

2.2 Les principaux modèles

Les principaux modèles de risques concurrents utilisés en fiabilité et en analyse de la maintenance sont les suivants:

- Le modèle à risques indépendants est la situation la plus simple où l'on considère Y et Z indépendants. Le modèle est alors complètement défini par les lois marginales de Y et Z. Comme nous considérons des MC et MP conditionnelles toutes deux liées à la dégradation du système, il n'est pas réaliste d'utiliser ce modèle bien que les équations qui en découlent soient très simples.

- Les modèles à hasards proportionnels consistent à supposer que le type (U) et l'instant (W) de la première maintenance sont indépendants. Pour ces modèles, la fonction ϕ est constante.

- Le modèle le plus employé en fiabilité est le modèle Delay-Time (Christer, 2002): $Y = A + B$ et $Z = A + C$, avec A, B et C, variables aléatoires mutuellement indépendantes. A correspond à une alerte émise avant la défaillance et B et C sont les temps d'attente après cette alerte avant d'effectuer une MP ou d'observer une défaillance. Quand B et C sont de lois exponentielles, ϕ est constante.

- Les modèles faisant l'hypothèse de signe aléatoire consistent à supposer que le type (U) de la première maintenance et l'instant (Z) de la potentielle défaillance sont indépendants. Pour ces modèles, la fonction ϕ est maximum à l'origine. Parmi ces modèles, on peut citer le modèle IPRA (Langseth et Lindqvist, 2003) pour lequel ϕ est décroissante.

- Le modèle Alert-Delay (Dijoux et Gaudoin, 2008): $Y = pZ + \mathcal{E}$, avec Z et $\mathcal{E}$ deux variables positives indépendantes. Le système émet un signal d'alerte à un instant proportionnel à la date de la potentielle défaillance. Il faut alors un délai $\mathcal{E}$ avant de pouvoir effectuer une MP. Si ce délai est trop long, on observe une panne suivie d'une MC. Lorsque Z et $\mathcal{E}$ suivent des lois exponentielles, la fonction ϕ est croissante.

3. Etude de l'estimateur non paramétrique de ϕ

3.1 Définition et propriétés

Dans une approche statistique, on étudie un système réparable soumis à une séquence de maintenances préventives et correctives. Cela consiste à observer une suite de durées inter-maintenance et leur type de maintenance associé, à savoir une séquence de k couples $(W_i, U_i)_{\{i\in[1,k]\}}$. Dans le cadre des risques concurrents classiques, on suppose les maintenances parfaites et donc que les $(W_i, U_i)_{\{i\geq 1\}}$ sont iid. L' estimateur de ϕ est alors:

$$\widehat{\phi_k}(w) = \begin{cases} \frac{\sum_{i=1}^{k} U_i \mathbb{1}_{\{W_i > w\}}}{\sum_{i=1}^{k} \mathbb{1}_{\{W_i > w\}}} & \text{si } \max_{i=1..k} W_i > w \\ \gamma(W_1, \ldots, U_k) \in [0,1] & \text{si } \max_{i=1..k} W_i \leq w \end{cases} \tag{1}$$

où $\gamma(W_1, \ldots, U_k)$ est une fonction déterministe du passé du processus. Nous avons prouvé le résultat suivant sur cet estimateur:

Théorème 1 *: $\widehat{\phi_k}$ est un estimateur asymptotiquement sans biais et convergent de ϕ. Plus précisément, le biais $\mathcal{B}[\widehat{\phi_k}(w)]$ vaut:*

$$\mathcal{B}[\widehat{\phi_k}(w)] = E[\widehat{\phi_k}(w)] - \phi(w) = (E[\gamma(W_1, \ldots, U_k) | \max_{i=1..k} W_i \leq w] - \phi(w))(1 - S_W(w))^k$$

Cet estimateur est sans biais lorsque $E[\gamma(W_1, \ldots, U_k) | \max_{i=1..k} W_i \leq w] = \phi(w)$. Il parait alors malheureusement difficile d'avoir un estimateur sans biais de ϕ en w. Toutefois, on peut proposer plusieurs choix pour $\gamma(w_1, \ldots, u_k)$:

- $\gamma(W_1, \ldots, U_k)$=0. Cela signifie qu'au-delà de $\max_{i=1..k} W_i$, on considère que toutes les maintenances sont correctives. Le biais est alors $-\phi(w)(1 - S_W(w))^k$.

- $\gamma(W_1, \ldots, U_k)$=1. Au-delà de $\max_{i=1..k} W_i$, on considère que toutes les maintenances sont préventives. Le biais est $\mathcal{B}[\widehat{\phi_k}(w)] = (1 - \phi(w))(1 - S_W(w))^k$.

- $\gamma(W_1, \ldots, U_k) = \frac{\sum_{i=1}^k U_i}{k}$. Cela revient à estimer $\phi(w)$ par la proportion de MP. On obtient alors $\mathcal{B}[\widehat{\phi_k}(w)] = (\phi(0) - \phi(w))(1 - S_W(w))^{k-1}$.

- $\gamma(W_1, \ldots, U_k) = U^*$, où U^* est le type de maintenance associé à la plus grande valeur de W_i. Il parait vraisemblable que pour ce choix de $\gamma(W_1, \ldots, U_k)$, l'estimation de $\phi(w)$ sera meilleure que les précédentes si ϕ est monotone.

3.2 Discussion

En comparant la forme de l'estimateur et celle de la fonction ϕ théorique pour plusieurs modèles, on peut déterminer quel modèle est le plus adapté pour des données. Si $\widehat{\phi_k}$ est maximum à l'origine, on peut privilégier un modèle de signe aléatoire. Si $\widehat{\phi_k}$ est pratiquement constante, on peut opter pour un modèle à hasards proportionnels. Si $\widehat{\phi_k}$ est croissante, on peut opter pour le modèle Alert-Delay.

Intéressons nous à un jeu de données provenant de données de maintenances pour deux compresseurs identiques provenant d'une usine Norsk-Hydro productrice d'ammoniac. Sur 21 années d'observation, 321 maintenances ont été effectuées, 189 MP et 132 MC. L'estimateur non paramétrique de ϕ est présenté sur la figure 1. On observe que $\widehat{\phi_k}$ fluctue entre les valeurs 0.6 et 0.7 et semble légèrement croître. La section suivante apporte des indications supplémentaires sur le modèle à privilégier.

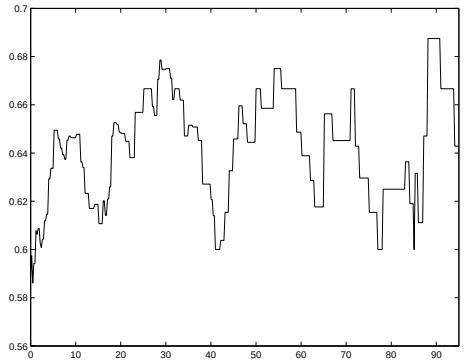

Figure 1: Estimation non paramétrique de ϕ

4. Test de monotonie de ϕ

La visualisation de $\widehat{\phi_k}$ donne des indications précieuses sur le choix d'un modèle. Dans le contexte de la maintenance, Dewan et al. (2004) ont proposé une procédure statistique basée sur le test de Kendall afin de tester si ϕ est constante. Ces tests sont basés sur les travaux de Deshpande et Sengupta (1995). Par la suite, nous proposons une procédure de tests non paramétriques, facilement implémentable, basée sur le test de Wilcoxon, toujours afin de tester si ϕ est constante.

Le principe est le suivant : si la fonction ϕ associée à un modèle est constante, cela signifie que W et U sont indépendants, ce qui correspond à un modèle à hasards proportionnels. En se plaçant sous l'hypothèse que ϕ est constante (H_0), supposons que l'on observe un échantillon de k maintenances $(W_i, U_i)_{\{1\leq i\leq k\}}$. Comme W et U sont indépendants, on peut considérer que les durées de MC observées $(W_i, U_i = 0)_{\{1\leq i\leq k\}}$ et les durées de MP observées $(W_i, U_i = 1)_{\{1\leq i\leq k\}}$ sont deux sous-échantillons de $(W_i)_{\{1\leq i\leq k\}}$. Soit $W_1^*, \ldots, W_k^*$ la statistique d'ordre de l'échantillon $W_1, \ldots, W_k$ et U_i^* le type de la maintenance associé à la durée W_i^*. Sous H_0, comme W et U sont indépendants, le rang d'un instant d'observation est indépendant du type de la maintenance. Ainsi, les deux sous-échantillons doivent être répartis de manière homogène dans l'échantillon complet. En revanche, si ϕ est croissante, les rangs de $(W_i, U_i = 1)_{\{1\leq i\leq k\}}$ doivent être plus élevés que $(W_i, U_i = 0)_{\{1\leq i\leq k\}}$. Le test de Wilcoxon permet justement de comparer deux échantillons indépendants. Il est basé sur la somme des rangs des observations du second échantillon dans l'échantillon complet. On aboutit à la propriété suivante:

Propriété 1 *Sous H_0 (ϕ constante), $T = \frac{2\sum_{i=1}^k iU_i^* - (k+1)\sum_{i=1}^k U_i}{\sqrt{(k+1)(\sum_{i=1}^k U_i)(k-\sum_{i=1}^k U_i)}}\sqrt{3} \xrightarrow{\mathcal{L}} \mathcal{N}(0,1)$*

où $\mathcal{N}(0,1)$ désigne la loi normale centrée réduite. En pratique, il est nécessaire d'observer au moins 8 MP et 8 MC afin de pouvoir faire une approximation normale. Autrement, on doit utiliser les tables de la loi de Wilcoxon. Sous l'approximation normale, on construit les tests en définissant leur région critique $\mathcal{W}$. On obtient :

Théorème 2 *Tests de monotonie de ϕ de seuil α :*

- *(i) Test de "$H_0 : \phi$ constante" contre "$H_1 : \overline{H_0}$": $\mathcal{W} = \left\{|T| > u_{\frac{\alpha}{2}}\right\}$.*
- *(ii) Test de "$H_2 : \phi$ décroissante" contre "$H_3 : \phi$ croissante": $\mathcal{W} = \{T > u_\alpha\}$.*
- *(iii) Test de "$H_3 : \phi$ croissante" contre "$H_2 : \phi$ décroissante": $\mathcal{W} = \{T < -u_\alpha\}$.*

  où u_α est le quantile d'ordre $1-\alpha$ de la loi normale centrée réduite.

Bien que construite de manière très différente, la statistique T_D proposée par Deshpande et Sengupta (1995) est asymptotiquement équivalente à la statistique T. On peut montrer en effet que $T_D = \sqrt{\frac{k(k+1)}{(k-1)^2}}\ T$.

Pour illustrer la démarche, on teste sur les données Norsk-Hydro de la figure 1 si la fonction ϕ est croissante en utilisant le test *(ii)*. On obtient comme statistique $T = 1.74$. La p-valeur correspondante est $\alpha_c = 4.1\%$. En particulier, en se fixant un seuil de 5%, on conclut que ϕ est croissante, ce qui est conforme à la visualisation de $\widehat{\phi}$. Le test confirme donc qu'un modèle pertinent pour ces données doit avoir une fonction ϕ croissante.

5. Conclusion et perspectives

Dans ce papier, nous avons présenter deux méthodes, basées sur l'étude de la fonction ϕ, afin d'avoir des critères de sélection de modèle de risques concurrents. En premier lieu, le comportement de l'estimateur non paramétrique de ϕ permet visuellement d'opter pour un modèle adapté. On a montré la convergence de cet estimateur. Dans un deuxième temps, on peut effectuer un test de monotonie de ϕ pour affiner le choix de modèle. Toutefois, ces méthodes restent très globales. Il serait intéressant de trouver une procédure qui puisse évaluer la pertinence d'un modèle de risque concurrent particulier, comme des statistiques d'adéquation d'un modèle pour un jeu de données. De plus, nous avons porté notre attention sur la fonction ϕ, mais d'autres fonctions estimables comme les fonctions de sous-survies conditionnelles peuvent être employées. Enfin, les méthodes proposées se placent dans le cadre où les maintenances sont parfaites, et il serait intéressant de généraliser ces critères en présence de maintenance imparfaite.

Bibliographie

[1] A. Christer (2002), A review of delay time analysis for modelling plant maintenance. In Osaki, S., editor, *Stochastic models in reliability and maintenance*, 89–124, Springer.

[2] R.M. Cooke, J. Paulsen (1994), Concepts for measuring maintenance performance and methods for analysing competing failure modes, *Proceedings of the ESREL Conference, La Baule*, 497–506.

[3] R.M. Cooke, T. Bedford (2002), Reliability databases in perspective, *IEEE Transactions on Reliability*, 51(3), 294-310.

[4] M.J. Crowder (2001), *Classical competing risks*, Chapman & Hall, London.

[5] J.V. Deshpande, D. Sengupta (1995), Testing the hypothesis of proportional hazards in two populations, *Biometrika*, 82(2), 251–61.

[6] I. Dewan, J.V. Deshpande, S.B. Kulathinal (2004), On Testing Dependence between Time to Failure and Cause of Failure via Conditional Probabilities, *Scandinavian Journal of Statistics*, 31, 79–91.

[7] Y. Dijoux, O. Gaudoin (2008), The alert-delay competing risks model for maintenance analysis, *Journal of Statistical Planning and Inference*, à paraître, disponible en ligne, doi: 10.1016/j.jspi.2008.05.021.

[8] H. Langseth, B.H. Lindqvist (2003), A maintenance model for components exposed to several failure mechanisms and imperfect repair. In Doksum, K. and Lindqvist, B.H., editors, *Mathematical and Statistical methods in Reliability*, Quality, Reliability and Engineering Statistics, 355–370. World Scientific Publishing Co.