

Estimation Bayésienne de la Durée de Fonctionnement des Véhicules de Chantier : Une Approche par Modèles Hiérarchiques

Stephen Cohen - ISUP 2025

Dans cet article, nous développons une approche bayésienne pour l'estimation de la durée de fonctionnement des véhicules de chantier dans un contexte assurantiel. Le modèle proposé prend en compte la structure séquentielle des périodes de fonctionnement et de maintenance, ainsi que l'influence des covariables environnementales. Notre contribution principale réside dans la construction d'un modèle hiérarchique combinant des lois de Weibull et exponentielles, dont les paramètres sont régis par des réseaux de neurones bayésiens.

- Estimation Bayésienne de la Durée de Fonctionnement des Véhicules de Chantier : Une Approche par Modèles Hiérarchiques
 - 1. Introduction
 - 1.1 Contexte
 - 1.2 Formalisation du Problème
 - 1.3 Objectifs
 - 2. État de l'Art
 - 2.1 Modèles de Survie en Maintenance Industrielle
 - 2.2 Approches Bayésiennes en Fiabilité
 - 2.3 Réseaux de Neurones Bayésiens pour la Prédiction de Durée de Vie
 - 3. Enrichissement du Modèle Mathématique
 - 3.1 Théorème de Représentation pour le Réseau Neuronale
 - 3.2 Propriétés du Modèle de Survie
 - 3.3 Estimateur de Bayes pour la Durée de Vie Résiduelle
 - 3.4 Bornes de Confiance Bayésiennes
 - 3.5 Analyse de Sensibilité et Robustesse
 - 3.6 Optimisation des Hyperparamètres
 - 3.7 Convergence de l'Algorithme MCMC
 - 4. Discussion des Implications Théoriques
 - 4.1 Propriétés Asymptotiques
 - 4.2 Complexité Computationnelle

- 4.3 Distributions A Priori
- 4.4 Vraisemblance et Censure
- 4.5 Extension au Cas Non-Stationnaire
- 5. Inférence Bayésienne
 - 5.1 Distribution A Posteriori
 - 5.2 Algorithme MCMC
 - 5.3 Estimateurs Bayésiens
- Références

1. Introduction

1.1 Contexte

Le problème de l'estimation de la durée de fonctionnement des véhicules de chantier présente des enjeux majeurs pour le secteur assurantiel. La complexité de cette estimation provient de la nature séquentielle des données et de la censure à droite inhérente aux observations.

1.2 Formalisation du Problème

Soit Ω un ensemble de véhicules professionnels indexés par $i \in 1, \dots, n$. Pour chaque véhicule i , nous observons une succession de périodes de fonctionnement et de maintenance :

$$(x_{i,1}, y_{i,1}, \delta_{i,1}), (\tilde{x}_i, 1, t_i, 1), \dots, (x_{i,m_i}, y_{i,m_i}, \delta_{i,m_i}), (\tilde{x}_i, m_i, t_i, m_i)$$

où :

$x_{i,k} \in \mathbb{R}^{d_1}$ représente les covariables du chantier $y_{i,k} \in \mathbb{R}_+$ est la durée de fonctionnement $\delta_{i,k} \in 0, 1$ est l'indicateur de censure $\tilde{x}_i \in \mathbb{R}^{d_2}$ décrit les conditions de maintenance $t_i \in \mathbb{R}_+$ est le temps de maintenance $m_i \in \mathbb{N}^*$ est le nombre de périodes observées

1.3 Objectifs

Notre objectif est triple :

- Développer un modèle bayésien robuste
- Quantifier l'incertitude des prédictions
- Fournir des recommandations assurantielles

2. État de l'Art

2.1 Modèles de Survie en Maintenance Industrielle

La modélisation de la durée de vie des équipements industriels a fait l'objet de nombreuses recherches. Cox et Oakes (1984) ont établi les fondements de l'analyse de survie moderne avec leur modèle à risques proportionnels. Dans le contexte spécifique de la maintenance industrielle, Jardine et al. (2006) ont développé des modèles de maintenance prédictive basés sur les processus de Markov.

Le modèle de Cox classique s'écrit :

$$h(t|x) = h_0(t) \exp(\beta^T x)$$

où $h_0(t)$ est le risque de base et $\beta^T x$ représente l'effet des covariables.

2.2 Approches Bayésiennes en Fiabilité

L'inférence bayésienne dans les modèles de fiabilité a été explorée par Ibrahim et al. (2001). Leur approche fondamentale consiste à modéliser la distribution a priori des paramètres de fiabilité par :

$$p(\theta) \propto \prod_{j=1}^p \text{IG}(\alpha_j, \beta_j)$$

où IG désigne la loi inverse-gamma.

2.3 Réseaux de Neurones Bayésiens pour la Prédiction de Durée de Vie

Neal (2012) a introduit l'utilisation des réseaux de neurones bayésiens pour la modélisation de la survie. Son approche combine :

- Une architecture feed-forward classique
- Des priors gaussiens sur les poids
- Une inférence par MCMC

3. Enrichissement du Modèle Mathématique

3.1 Théorème de Représentation pour le Réseau Neuronal

Théorème 1. *Pour notre architecture neuronale g_1 , avec des activations ReLU et K neurones dans la couche cachée, nous avons la représentation :*

$$g_1(x, t, \theta_1) = \exp \left(\sum_{k=1}^K w_k^{(2)} \max(0, w_k^{(1)T} [x, t] + b_k^{(1)}) + b^{(2)} \right)$$

où $[x, t]$ désigne la concaténation des vecteurs x et t .

Démonstration.

1. Soit $h_k^{(1)} = \max(0, w_k^{(1)T} [x, t] + b_k^{(1)})$ la sortie du k-ème neurone de la première couche
2. La sortie de la deuxième couche est $\sum_{k=1}^K w_k^{(2)} h_k^{(1)} + b^{(2)}$
3. L'application de la transformation exponentielle donne le résultat.

3.2 Propriétés du Modèle de Survie

Théorème 2. Pour notre modèle de Weibull modifié, la fonction de survie conditionnelle s'écrit :

$$S(y|x, t, \theta_1, \beta) = \exp \left(- \left(\frac{y}{g_1(x, t, \theta_1)} \right)^\beta \right)$$

et possède la propriété de hasard proportionnel vis-à-vis des covariables.

Démonstration.

1. Par définition, $S(y|x, t, \theta_1, \beta) = P(Y > y|x, t, \theta_1, \beta)$
2. Pour la loi de Weibull :

$$\begin{aligned} S(y|x, t, \theta_1, \beta) &= 1 - F(y|x, t, \theta_1, \beta) \\ &= 1 - \left(1 - \exp \left(- \left(\frac{y}{g_1(x, t, \theta_1)} \right)^\beta \right) \right) \end{aligned}$$

3. La fonction de hasard est :

$$\begin{aligned} h(y|x, t, \theta_1, \beta) &= \frac{f(y|x, t, \theta_1, \beta)}{S(y|x, t, \theta_1, \beta)} \\ &= \frac{\beta}{\eta} \left(\frac{y}{\eta} \right)^{\beta-1} \end{aligned}$$

où $\eta = g_1(x, t, \theta_1)$

3.3 Estimateur de Bayes pour la Durée de Vie Résiduelle

Pour la prédiction de la durée de vie résiduelle, nous définissons :

Définition 1. L'estimateur de Bayes de la durée de vie résiduelle sous la fonction de perte quadratique est :

$$\hat{Y}_{i,k+1} = \mathbb{E}[Y_{i,k+1} | x_{i,k+1}, t_{i,k}, \mathcal{D}]$$

Théorème 3. Pour notre modèle, cet estimateur s'écrit :

$$\hat{Y}_{i,k+1} = \int_{\Theta} g_1(x_{i,k+1}, t_{i,k}, \theta_1) \Gamma(1 + \frac{1}{\beta}) p(\theta_1, \beta | \mathcal{D}) d\theta_1 d\beta$$

Démonstration.

1. Par la loi de l'espérance totale :

$$\hat{Y}_{i,k+1} = \int_{\Theta} \mathbb{E}[Y_{i,k+1} | \theta_1, \beta, x_{i,k+1}, t_{i,k}] p(\theta_1, \beta | \mathcal{D}) d\theta_1 d\beta$$

2. Pour une loi de Weibull de paramètres (η, β) :

$$\mathbb{E}[Y] = \eta \Gamma(1 + \frac{1}{\beta})$$

3. En substituant $\eta = g_1(x_{i,k+1}, t_{i,k}, \theta_1)$, on obtient le résultat.

3.4 Bornes de Confiance Bayésiennes

Théorème 4. Les intervalles de crédibilité à $(1 - \alpha)100\%$ pour la durée de vie résiduelle sont donnés par :

$$[Q_{\alpha/2}(Y_{i,k+1} | \mathcal{D}), Q_{1-\alpha/2}(Y_{i,k+1} | \mathcal{D})]$$

où Q_p désigne le quantile d'ordre p de la distribution prédictive postérieure.

3.5 Analyse de Sensibilité et Robustesse

Théorème 6. (Robustesse du modèle) Sous les hypothèses de régularité du réseau neuronal et pour des perturbations bornées δ des covariables, on a :

$$\sup_{x,t} |g_1(x + \delta, t, \theta_1) - g_1(x, t, \theta_1)| \leq L \|\delta\|$$

où L est une constante de Lipschitz qui dépend de la structure du réseau.

Preuve : La démonstration utilise les propriétés de régularité des activations ReLU et la structure en couches du réseau.

3.6 Optimisation des Hyperparamètres

Le choix optimal de la structure du réseau (nombre de couches et de neurones) peut être guidé par le critère d'information bayésien (BIC) :

$$\text{BIC} = -2 \ln(\mathcal{L}) + k \ln(n)$$

où :

- \mathcal{L} est la vraisemblance maximale
- k est le nombre de paramètres
- n est la taille de l'échantillon

3.7 Convergence de l'Algorithme MCMC

Théorème 7. (Convergence de l'échantillonneur) *Pour l'algorithme HMC-NUTS proposé, sous des conditions de régularité standards :*

1. La chaîne est géométriquement ergodique
2. L'erreur d'estimation décroît en $O(1/\sqrt{N})$ où N est le nombre d'itérations

4. Discussion des Implications Théoriques

4.1 Propriétés Asymptotiques

Le théorème suivant établit la consistance de notre estimateur :

Théorème 5. *Sous des conditions de régularité standards (continuité et bornitude des réseaux neuronaux), notre estimateur bayésien est consistant au sens où :*

$$\hat{Y}_{i,k+1} \xrightarrow{p} Y_{i,k+1}^*$$

où $Y_{i,k+1}^*$ est la vraie valeur de la durée de vie résiduelle.

La démonstration complète utilise les résultats de Ghosal et van der Vaart (2017) sur la consistance bayésienne.

4.2 Complexité Computationnelle

L'inférence bayésienne dans notre modèle nécessite :

1. $O(NK)$ opérations pour l'évaluation du réseau neuronal
2. $O(L)$ itérations MCMC
3. Complexité totale en $O(NKL)$ où :

- N est la taille de l'échantillon
- K est le nombre de neurones
- L est le nombre d'itérations MCMC

4.3 Distributions A Priori

Pour les poids des réseaux neuronaux :

$$W_j^{(1)} \sim \mathcal{N}(0, \tau_j^2 I)$$

$$W_j^{(2)} \sim \mathcal{N}(0, \nu_j^2 I)$$

Pour les hyperparamètres :

$$\tau_j^2 \sim \mathcal{IG}(\alpha_\tau, \beta_\tau)$$

$$\nu_j^2 \sim \mathcal{IG}(\alpha_\nu, \beta_\nu)$$

Pour β :

$$\beta \sim \text{TBeta}(a_\beta, b_\beta, 1, 3)$$

4.4 Vraisemblance et Censure

La vraisemblance complète :

$$\mathcal{L}(\theta_1, \theta_2, \beta | \mathcal{D}) = \prod_{i=1}^n \prod_{k=1}^{m_i} \mathcal{L}Y(y_i, k | \theta_1, \beta) \mathcal{L}T(t_i, k | \theta_2)$$

Pour les données censurées ($\delta_{i,k} = 0$) :

$$\mathcal{L}Y(y_i, k | \theta_1, \beta) = \exp \left(- \left(\frac{y_{i,k}}{\eta_{i,k}} \right)^\beta \right)$$

Pour les données non censurées ($\delta_{i,k} = 1$) :

$$\mathcal{L}Y(y_i, k | \theta_1, \beta) = f_W(y_{i,k} | \eta_{i,k}, \beta)$$

4.5 Extension au Cas Non-Stationnaire

Le modèle peut être étendu pour prendre en compte l'évolution temporelle des paramètres :

$$g_1(x, t, \theta_1, \tau) = g_1(x, t, \theta_1) \exp(h(\tau))$$

où $h(\tau)$ est une fonction de tendance temporelle et τ le temps calendaire.

5. Inférence Bayésienne

5.1 Distribution A Posteriori

$$p(\theta_1, \theta_2, \beta | \mathcal{D}) \propto \mathcal{L}(\theta_1, \theta_2, \beta | \mathcal{D}) p(\theta_1) p(\theta_2) p(\beta)$$

5.2 Algorithme MCMC

Nous utilisons l'algorithme HMC-NUTS pour :

- Gérer les corrélations entre paramètres
- S'adapter à la géométrie de l'espace
- Échantillonner efficacement

5.3 Estimateurs Bayésiens

Théorème 2. (Estimateur de la durée de vie résiduelle) L'estimateur de Bayes sous perte quadratique est :

$$\hat{Y}_{i,k+1} = \int \Theta g_1(x_{i,k+1}, t_{i,k}, \theta_1) \Gamma(1 + \frac{1}{\beta}) p(\theta_1, \beta | \mathcal{D}) d\theta_1 d\beta$$

Références

1. Cox, D. R., & Oakes, D. (1984). Analysis of survival data. Chapman and Hall/CRC.
2. Jardine, A. K., Lin, D., & Banjevic, D. (2006). A review on machinery diagnostics and prognostics implementing condition-based maintenance. Mechanical systems and signal processing, 20(7), 1483-1510.
3. Ibrahim, J. G., Chen, M. H., & Sinha, D. (2001). Bayesian survival analysis. Springer Science & Business Media.
4. Neal, R. M. (2012). Bayesian learning for neural networks (Vol. 118). Springer Science & Business Media.
5. Ghosal, S., & van der Vaart, A. (2017). Fundamentals of nonparametric Bayesian inference (Vol. 44). Cambridge University Press.