Modèles de Détection de Rupture

Le Cnam- EICNAM

9 mars 2020

Sommaire

Modèle

Soit $X_1, \ldots, X_n$ des variables aléatoires indépendantes. On suppose qu'il existe un instant k $(1 < k < n-1)$, appelé l'instant de Rupture tel que :

Modèle

Soit $X_1, \ldots, X_n$ des variables aléatoires indépendantes. On suppose qu'il existe un instant k $(1 < k < n-1)$, appelé l'instant de Rupture tel que :

Cas discret

$$\begin{cases} X_1, \ldots, X_k \ \text{ suivent la loi } \ P_{\theta_1} \\ X_{k+1}, \ldots, X_n \ \text{ suivent la loi } \ P_{\theta_2} \end{cases}$$

Modèle

Soit $X_1, \ldots, X_n$ des variables aléatoires indépendantes. On suppose qu'il existe un instant k $(1 < k < n-1)$, appelé l'instant de Rupture tel que :

Cas discret

$$\begin{cases} X_1, \ldots, X_k & \text{suivent la loi} \quad P_{\theta_1} \\ X_{k+1}, \ldots, X_n & \text{suivent la loi} \quad P_{\theta_2} \end{cases}$$

Cas continu

$$\begin{cases} X_1, \ldots, X_k & \text{suivent la loi} \quad f_1(x, \theta_1) \\ X_{k+1}, \ldots, X_n & \text{suivent la loi} \quad f_2(x, \theta_2) \end{cases}$$

Fonction de vraisemblance

Notons L la vraisemblance associée au modèle et ℓ la fonction Log-vraisemblance.

Fonction de vraisemblance

Notons L la vraisemblance associée au modèle et ℓ la fonction Log-vraisemblance.

Cas discret

Fonction de vraisemblance

Notons L la vraisemblance associée au modèle et ℓ la fonction Log-vraisemblance.

Cas discret

$$\begin{cases} L(\underline{x}, \theta_1, \theta_2, k) = \prod_{i=1}^{k} P_{\theta_1}(X_i = x_1) \prod_{i=k+1}^{n} P_{\theta_2}(X_i = x_1) \\ \ell(\underline{x}, \theta_1, \theta_2, k) = \sum_{i=1}^{k} \log\left(P_{\theta_1}(X_i = x_i)\right) + \sum_{i=k+1}^{n} \log\left(P_{\theta_2}(X_i = \right. \end{cases}$$

Notons L la vraisemblance associée au modèle et ℓ la fonction Log-vraisemblance.

Cas discret

$$\begin{cases} L(\underline{x}, \theta_1, \theta_2, k) = \prod_{i=1}^{k} P_{\theta_1}(X_i = x_1) \prod_{i=k+1}^{n} P_{\theta_2}(X_i = x_1) \\ \ell(\underline{x}, \theta_1, \theta_2, k) = \sum_{i=1}^{k} \log\left(P_{\theta_1}(X_i = x_i)\right) + \sum_{i=k+1}^{n} \log\left(P_{\theta_2}(X_i = \right. \end{cases}$$

Cas continu

Fonction de vraisemblance

Notons L la vraisemblance associée au modèle et ℓ la fonction Log-vraisemblance.

Cas discret

$$\begin{cases} L(\underline{x}, \theta_1, \theta_2, k) = \prod_{i=1}^{k} P_{\theta_1}(X_i = x_1) \prod_{i=k+1}^{n} P_{\theta_2}(X_i = x_1) \\ \ell(\underline{x}, \theta_1, \theta_2, k) = \sum_{i=1}^{k} \log\left(P_{\theta_1}(X_i = x_i)\right) + \sum_{i=k+1}^{n} \log\left(P_{\theta_2}(X_i = \right. \end{cases}$$

Cas continu

$$\begin{cases} L(\underline{x}, \theta_1, \theta_2, k) = \prod_{i=1}^{k} f_1(x_i, \theta_1) \prod_{i=k+1}^{n} f_2(x_i, \theta_2) \\ \ell(\underline{x}, \theta_1, \theta_2, k) = \sum_{i=1}^{k} \log\left(f_1(x_i, \theta_1)\right) + \sum_{i=k+1}^{n} \log\left(f_2(x_i, \theta_2)\right) \end{cases}$$

Estimation des paramètres

Les estimateurs sont obtenus par la méthode du maximum de vraisemblance :

Estimation des paramètres

Les estimateurs sont obtenus par la méthode du maximum de vraisemblance :

$$(\hat{\theta}_1(k), \hat{\theta}_2(k)) = \mathrm{Arg} \sup_{(\theta_1,\theta_2)} \ell(\underline{x}, \theta_1, \theta_2, k) \tag{1}$$

Estimation des paramètres

Les estimateurs sont obtenus par la méthode du maximum de vraisemblance :

$$(\hat{\theta}_1(k), \hat{\theta}_2(k)) = \mathrm{Arg} \sup_{(\theta_1, \theta_2)} \ell(\underline{x}, \theta_1, \theta_2, k) \tag{1}$$

estimation de l'intant de Rupture

Notons $\Lambda(k) = \ell(\underline{x}, \hat{\theta}_1(k), \hat{\theta}_2(k), k)$. $\hat{k}$ l'estimateur du maximum de vraisemblance de k est déterminé par :

$$\hat{k} = \mathrm{Arg} \max_{k} \Lambda(k) \tag{2}$$

Modèle de Rupture pour la loi de Pareto

Soit $f(x,a) = \dfrac{a}{x^{a+1}} \mathbb{1}_{[1,\infty[}(x) = a\, e^{-(a+1)\log x} \mathbb{1}_{[1,\infty[}(x)$ où $a > 0$.

Soit $X_1, \ldots, X_n$ des variables aléatoires indépendantes. On suppose qu'il existe un instant k $(1 < k < n-1)$, appelé l'instant de Rupture tel que :

Modèle de Rupture pour la loi de Pareto

Soit $f(x,a) = \dfrac{a}{x^{a+1}} \mathbb{1}_{[1,\infty[}(x) = a\, e^{-(a+1)\log x} \mathbb{1}_{[1,\infty[}(x)$ où $a > 0$.

Soit $X_1, \ldots, X_n$ des variables aléatoires indépendantes. On suppose qu'il existe un instant k $(1 < k < n-1)$, appelé l'instant de Rupture tel que :

Modèle de Rupture pour la loi de Pareto

$$\begin{cases} X_1, \ldots, X_k & \text{suivent la loi} \quad f(x_i, a_1) \\ X_{k+1}, \ldots, X_n & \text{suivent la loi} \quad f(x_i, a_2) \end{cases}$$

Modèle de Rupture pour la loi de Pareto

vraisemblance et la fonction Log-vraisemblance associée au modèle

$$
\begin{cases}
L(\underline{x}, a_1, a_2, k) = \prod_{i=1}^{k} a_1 \, e^{-(a_1+1)\log x_i} \prod_{i=k+1}^{n} a_2 \, e^{-(a_2+1)\log x_i} \\
\ell(\underline{x}, a_1, a_2, k) = \sum_{i=1}^{k} (\log a_1 - (a_1+1)\log x_i) \\
\quad + \sum_{i=k+1}^{n} (\log a_2 - (a_2+1)\log x_i) \\
\quad = k \log a_1 - (a_1+1) \sum_{i=1}^{k} \log x_i \\
\quad +(n-k)\log a_2 - (a_2+1) \sum_{i=k+1}^{n} \log x_i
\end{cases}
$$

Modèle de Rupture pour la loi de Pareto

estimateurs des paramètres

$$\begin{cases} \hat{a}_1 = \hat{a}_1(k) = \dfrac{k}{\displaystyle\sum_{i=1}^{k} \log X_i} \\ \hat{a}_2 = \hat{a}_2(k) = \dfrac{n-k}{\displaystyle\sum_{i=k+1}^{n} \log X_i} \end{cases} \tag{3}$$

Modèle de Rupture pour la loi de Pareto

estimateurs des paramètres

$$\begin{cases} \hat{a}_1 = \hat{a}_1(k) = \dfrac{k}{\displaystyle\sum_{i=1}^{k} \log X_i} \\ \hat{a}_2 = \hat{a}_2(k) = \dfrac{n-k}{\displaystyle\sum_{i=k+1}^{n} \log X_i} \end{cases} \tag{3}$$

Notons $\Lambda(k) = \ell(\underline{x}, \hat{a}_1(k), \hat{a}_2(k), k)$. $\hat{k}$ l'estimateur du maximum de vraisemblance de k est déterminé par :

$$\hat{k} = \mathrm{Arg}\max_{k} \Lambda(k) \tag{4}$$

estimateur de l'instant de rupture

$$\begin{aligned}
\Lambda(k) &= k \log \hat{a}_1(k) - (\hat{a}_1(k)+1) \sum_{i=1}^{k} \log X_i \\
&+ (n-k) \log \hat{a}_2(k) - (\hat{a}_2(k)+1) \sum_{i=k+1}^{n} \log X_i \\
&= k \log \left(\frac{k}{\sum_{i=1}^{k} \log X_i} \right) \\
&+ (n-k) \log \left(\frac{n-k}{\sum_{i=k+1}^{n} \log X_i} \right) - n - \sum_{i=1}^{n} \log X_i
\end{aligned} \tag{5}$$

Modèle de Rupture pour la loi de Pareto

Remarque pour la programmation

En programmation, pour l'expression de $\Lambda(k)$, on tient compte des termes qui font intervenir k :

$$\Lambda(k) = k \log\left(\frac{k}{\sum_{i=1}^{k} \log X_i}\right) + (n-k)\log\left(\frac{n-k}{\sum_{i=k+1}^{n} \log X_i}\right) \quad (6)$$