

**Sommaire**

---

1	Introduction . . . . .	61
2	Généralités . . . . .	62
3	Test de Page-Hinkley . . . . .	63
3.1	Cas de moyennes connues . . . . .	63
3.2	Cas de moyenne inconnu après rupture . . . . .	65
3.3	L'algorithme général du test de Page-Hinkley . . . . .	66
3.4	Exemple d'implémentation et mise en oeuvre du test de Page-Hinkley . . . . .	67

---

## 1 Introduction

L'analyse des résidu présent un étape primordiale qui viens après la génération des résidu et le diagnostic pour l'évaluation des résultats obtenus par les méthodes proposées précédemment. Elle est relatif à la théorie de décision, particulièrement aux test d'hypothèses. L'objectif de l'analyse des résidus ou la décision est d'effectuer un choix entre deux hypothèses : appelons  $H_0$  l'hypothèse associé au bon fonctionnement du système ou le fonctionnement normal, et  $H_1$  est l'hypothèse associé au mauvaise fonctionnement du système ou le fonctionnement défaillant. Ce choix est réalisé à partir d'une reconnaissance logique, grâce à un test qui associe pour chaque résidu  $i$  l'hypothèse  $H_0^i$  ou  $H_1^i$ . Ce test détecte un changement de moyenne ou de variance du résidu, qui révèle l'influence d'une défaillance. A partir des hypothèses effectuées sur les résidus, la procédure de localisation doit déterminer l'état de fonctionnement du système. La qualité du résultat final est fortement liée aux performances de l'algorithme de détection.

Généralement, deux principaux méthodes d'analyse des résidus, méthodes déterministes et stochastiques.

**Méthodes déterministes** Ce type des méthodes permettent d'utiliser des stratégies d'analyses et d'évaluation de la norme des résidus. Dans ce cas, le défaut affecte directement les valeurs des résidus. Ici, si les résidus présentent des valeurs à l'environnement de zéro donc le système est en fonctionnement normal, sinon le systèmes est en fonctionnement anormal. Pour cela, on peut déduire :

$$\begin{cases} H_0(0, t) & \text{si } \|r(t)\| \approx 0 \\ H_1(f_j, t_j) & \text{si } \|r(t_{\geq t_j})\| \neq 0 \end{cases}$$

avec  $\|r(t)\|$  est la norme des résidu,  $f_j$  défaut à l'instant  $t_j$ . Plusieurs méthodes existent dans cette catégorie tel que : méthode de retour à la norme, méthode de fonction de seuil, méthode de fonction d'évaluation.

**Méthodes stochastiques** Cette catégorie est basé sur les méthodes statistiques. Dans ce type les résidus présentent une nature aléatoire d'une probabilité  $p$ . En effet, la modification de la probabilité  $p$  ou le changement de caractéristiques statiques présente le changement de fonctionnement du système. L'un des méthodes les plus connues et les plus efficace dans

cette catégorie est la méthode du test de Page-Hinkley.

Dans ce chapitre on s'intéresse à la deuxième catégorie des méthodes stochastiques, plus précisément la méthode du test de Page-Hinkley.

## 2 Généralités

Généralement les signaux à détecter pour un saut de moyenne ou d'écart-type sont considérés comme gaussiens. Pour lequel, la loi de probabilité d'un signal  $r$  ayant une distribution gaussienne est donnée par :

$$R(r) = \frac{1}{\sqrt{2\pi\sigma}} \exp\left(\frac{-1}{2}\left(\frac{r-\mu}{\sigma}\right)^2\right) \quad (6.1)$$

avec  $\mu = Esp(r)$  et  $\sigma^2 = Var(r)$ . Chacun de ces paramètres est capable d'être influencés par une défaillance. Leur valeur peut varier aussi brusquement que la valeur du paramètre défaillant.

Pour  $K = 1 \dots n$ , le signal des résidus  $r$  décrit à chaque instant par :

$$r_k = \mu_k + \varepsilon_k \quad (6.2)$$

Avec  $\varepsilon$  représente un bruit gaussien centré de variance  $\sigma_k^2$  et  $\mu_k$  la valeur moyenne des résidus  $r_k$ . Les valeurs d'écart type et de moyenne des résidus  $r_k$  peuvent s'exprimer par :

$$\mu_0 = \frac{1}{n} \sum_{k=1}^n r_k \quad (6.3)$$

$$\sigma_0^2 = \frac{1}{n-1} \sum_{k=1}^n (r_k - \mu_0)^2 \quad (6.4)$$

On peut distinguer deux classes de changement, changement de moyenne et changement de variance. Pour l'instant  $K$ , le changement de moyenne des résidus  $r$  est décrit par le modèle suivant :

$$\mu_k = \begin{cases} \mu_0 & \text{si } k \leq K - 1 \\ \mu_1 & \text{si } k \geq K \end{cases} \quad (6.5)$$

En revanche, le changement de la variation des résidus  $r$  à l'instant  $K$  est exprimé par :

$$\varepsilon_k = \begin{cases} \sigma_0^2 & \text{si } k \leq K - 1 \\ \sigma_1^2 & \text{si } k \geq K \end{cases} \quad (6.6)$$

Le problème consiste à déterminer, à l'instant  $n$ , l'instant de changement  $K$  et la valeur des moyennes  $\mu_0$  et  $\mu_1$  avant et après changement.

Le signal est caractérisé par la loi normale notée  $N(\mu, \sigma^2)$ , et on se place dans le cas où le signal **conserve sa variance  $\sigma^2$  mais** passe d'une moyenne  $\mu_0$  à une moyenne  $\mu_1$  à l'instant  $K$ .

### 3 Test de Page-Hinkley

Le test de Page-Hinkley est un test statistique pour la détection du bordure, il a été introduit par Page(1954) comme une solution pour surveillance du comportement d'un processus industriel afin de surveiller la sortie du processus et détecter un éventuel changement ou une dérivation dans la sortie. L'objectif de cette méthode est d'étudier et **détecter** l'instant de l'apparition d'un **saut de moyenne**, en connaissant à l'avance la moyenne et la variance du signal.

#### 3.1 Cas de moyennes connues

Dans le cas de la moyenne connue, l'amplitude de la défaillance  $V$  supposée connue par contre pour l'instant de rupture  $K$  est inconnue.

La détection de ce type de rapture consiste à tester entre l'hypothèse de fonctionnement normal  $H_0$  et l'hypothèse de changement ou de défaillance  $H_1$  qui présentent les valeurs moyenne des résidus suivantes :

$$H_0 : \mu_k = \mu_0 \quad \text{Pour } k = 1, \dots, n \quad (6.7)$$

$$H_1 : \mu_k = \begin{cases} \mu_0 & \text{Pour } k = 1, \dots, K-1 \\ \mu_1 & \text{Pour } k = K, \dots, n \end{cases} \quad (6.8)$$

Le rapport de vraisemblance entre les deux hypothèses est défini par :

$$\prod_{k=K}^n \frac{R_1(r_k)}{R_0(r_k)} \quad (6.9)$$

Avec  $R$  est la probabilité gaussienne de moyenne  $\mu_{i,(i=,0)}$  et de variance  $\sigma_{i,(i=,0)} = \sigma^2$ . En utilisant la formulation 6.1, on remplaçant  $R_0$  et  $R_1$  par leur valeurs, 6.7 peut s'exprimer par :

$$\begin{aligned} \Lambda(r) &= \prod_{k=K}^n \exp \left\{ -\frac{1}{2\sigma^2} [(r_k - \mu_1)^2 - (r_k - \mu_0)^2] \right\} \\ &= \exp \left\{ -\frac{1}{2\sigma^2} \sum_{k=K}^n [(r_k - \mu_1)^2 - (r_k - \mu_0)^2] \right\} \end{aligned}$$

Le logarithme du rapport de vraisemblance est alors est :

$$\begin{aligned} \log \Lambda(r) &= \frac{\mu_1 - \mu_0}{\sigma^2} \sum_{k=K}^n \left( r_k - \frac{\mu_1 + \mu_0}{2} \right) \\ &= \frac{1}{\sigma^2} S_K^n(\mu_0, V) \end{aligned}$$

avec  $S_i^j(\mu, V) = V \sum_{k=i}^j (r_k - \mu - \frac{V}{2})$  et  $V = \mu_1 - \mu_0$  est l'amplitude du saut. En remplaçant l'instant  $K$  de changement par son estimation du maximum de vraisemblance sous  $H_1$  :

$$\hat{K} = \arg_{1 \leq K \leq n} \max \left[ \prod_{k=0}^{K-1} R_0(r_k) \prod_{k=0}^{K-1} R_1(r_k) \right] = \arg_{1 \leq K \leq n} \max S_K^n(\mu_0, V) \quad (6.10)$$

L'indicateur de changement est défini par :

$$g_n = \log \Lambda(r)(\hat{K}_n) = \max_K S_K^n(\mu_0, V) \quad (6.11)$$

à comparer à un seuil  $\lambda$ . On décide  $H_1$  et on arrête le test si  $g_n$  dépasse  $\lambda$ , sinon l'hypothèse  $H_0$  est conservée, et on continue le test jusqu'à décider  $H_1$ . On peut encore écrire ce détecteur de la manière suivante :

$$g_n = S_K^n(\mu_0, V) - \min_{1 \leq k \leq K} S_1^k(\mu_0, V) > \lambda \quad (6.12)$$

l'estimée de l'instant de rupture étant alors fournie par le dernier indice pour lequel la valeur minimum de  $S_1^k(\mu_0, V)$  a été atteinte.

### 3.2 Cas de moyenne inconnu après rupture

On se place désormais dans le cas le plus proche de la réalité : on ne connaît pas l'amplitude du saut  $V$ , c'est à dire que l'on ne connaît plus la moyenne  $\mu_1$  après le changement. Une autre approche consiste à exécuter en parallèle la règle d'arrêt de Page-Hinkley correspondant à un saut de moyenne minimum choisi  $V_m$  à détecter dans deux directions possibles : saut de moyenne positif ou négatif.

Les règles d'arrêt correspondantes sont les suivantes :

$$\text{Saut négative : } \begin{cases} T_0 = & 0 \\ T_n = & \sum_{k=1}^n (r_k - \mu_0 + \frac{V_m}{2}); \quad n \geq 1 \\ M_n = & \max_{0 \leq k \leq n} T_k \\ \text{détection si} & M - T_n > \lambda \end{cases}$$

$$\text{Saut positive : } \begin{cases} U_0 = & 0 \\ U_n = & \sum_{k=1}^n (r_k - \mu_0 + \frac{V_m}{2}); \quad n \geq 1 \\ m_n = & \min_{0 \leq k \leq n} U_k \\ \text{détection si} & M - U_n > \lambda \end{cases}$$

La décision prise est la première règle qui stoppe. Le temps du saut  $K$  est l'instant correspondant au dernier extremum avant la détection.

Une seconde approche consiste à remplacer l'amplitude du saut par son estimation au sens du maximum de vraisemblance :

$$H_1 : \max_{0 \leq k \leq n} \max_V S_K^n(\mu_0, V) > \lambda \quad (6.13)$$

Donc on a :

$$\hat{V}_n = \arg \max_V S_K^n(\mu_0, V) = \frac{1}{n - K + 1} \sum_{k=K}^n (r_k - \mu_0) \quad (6.14)$$

Ainsi, la double maximisation 6.14 est réduite à une seule maximisation. Pour ce test, deux paramètres sont à faire varier : l'amplitude de défaillance minimum à détecter  $V_m$  et le seuil  $\lambda$ . L'amplitude de panne minimum à détecter  $V_m$  peut être exprimée en nombre d'écart-type du signal  $\delta$ , soit :  $V_m = \delta \sigma_0$ . La détection est optimale lorsque l'amplitude de la défaillance estimée  $V_m$  est égale à l'amplitude de la défaillance  $V$ .

### 3.3 L'algorithme général du test de Page-Hinkley

L'algorithme général de la méthode est donné comme suite :

**Initialisation** contiens :

- Compte tenu du critère d'alarme  $\lambda$
- Toutes les données de résidus
- Le critère minimum de détection  $\delta$
- La moyenne du signal nominal

**processus** contiens :

- Nouvelle moyenne des données déterminé :

$$m\mu_0 = \frac{1}{n} \sum_{k=1}^n r_s$$

- Procéder au calcul de la somme cumulée :  $S_i^j(\mu, V) = V \sum_{k=i}^j (r_k - \mu - \frac{V}{2})$

- Extraire la valeur maximale de  $M_n = \max_{1 \leq k \leq K} S_1^k(\mu_0, V)$

- Comparaison avec le critère de détection
- Si  $M_n - S_1^k(\mu, V) \geq \lambda$  : détection positive
- Si  $M_n - S_1^k(\mu, V) \leq \lambda$  : détection négative

**Conclusion** contiens :

- Si la détection de front est vraie : arrêtez le système.
- Si la détection de front est fausse : lancez le processus pour les nouvelles données.

### 3.4 Exemple d'implémentation et mise en oeuvre du test de Page-Hinkley

Considérant une évolution d'un signal  $M - T_n$  présentée dans la figure 6.1. Le seuil de détection  $\lambda$  a été calculé en fixant la valeur minimale de l'amplitude du saut à détecter à  $V_m$ , ce qui correspond donc à la valeur **vraie**. Une surestimation de cette amplitude entraîne un seuil de détection plus important, augmentant ainsi le retard à la détection. Cependant, l'estimation de l'instant de la rupture n'en est pas affectée.

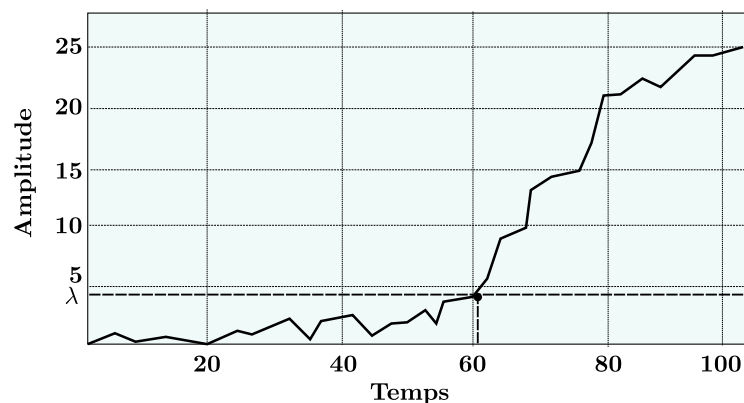


FIGURE 6.1 – Evolution de  $M - T_n$  et seuil

Le saut moyenne a été détecté à l'instant 61 et l'estimation de l'instant du changement est égale à 50 qui correspond au dernier instant pour lequel la valeur minimal de la somme cumulée  $T_n$  a été atteinte (Figure 6.2).

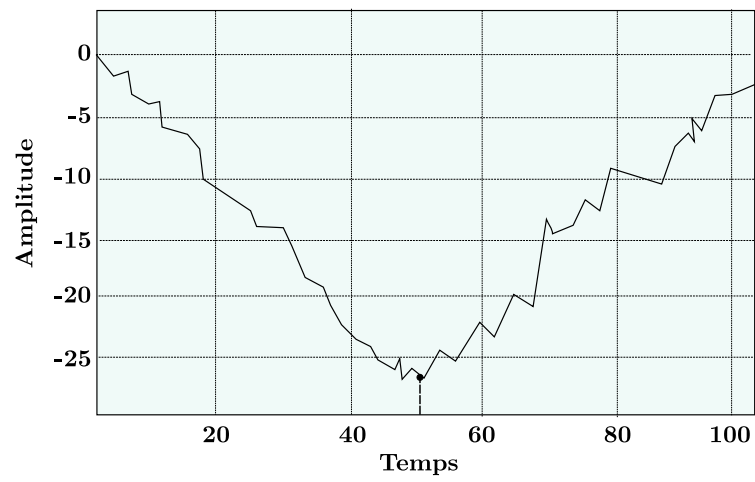


FIGURE 6.2 – Evolution de  $T_n$  et seuil