

# REVUE DE STATISTIQUE APPLIQUÉE

M. BLANCHET

M. BOUZID

R. DELHOME

J.-B. OUAKAM

M. PIETRI

J. BERNIER

## **Phénomènes de rupture dans une série chronologique**

*Revue de statistique appliquée*, tome 19, n° 4 (1971), p. 5-22

[http://www.numdam.org/item?id=RSA\\_1971\\_\\_19\\_4\\_5\\_0](http://www.numdam.org/item?id=RSA_1971__19_4_5_0)

© Société française de statistique, 1971, tous droits réservés.

L'accès aux archives de la revue « *Revue de statistique appliquée* » (<http://www.sfds.asso.fr/publicat/rsa.htm>) implique l'accord avec les conditions générales d'utilisation (<http://www.numdam.org/conditions>). Toute utilisation commerciale ou impression systématique est constitutive d'une infraction pénale. Toute copie ou impression de ce fichier doit contenir la présente mention de copyright.

NUMDAM

Article numérisé dans le cadre du programme  
Numérisation de documents anciens mathématiques  
<http://www.numdam.org/>

# PHÉNOMÈNES DE RUPTURE DANS UNE SÉRIE CHRONOLOGIQUE

M. BLANCHET, M. BOUZID, R. DELHOME, J.-B. OUKAM et M. PIETRI

sous la direction de J. BERNIER  
Professeur à l'I.S.U.P.

## INTRODUCTION

Il n'est pas rare d'observer des phénomènes de rupture dans la tendance d'une série chronologique par ailleurs homogène. Citons par exemple le dérèglement d'une machine dans une chaîne de fabrication, la brusque variation du trend d'une série économique.

Ces phénomènes sont souvent étudiés dans le cadre de la statistique descriptive (moyennes mobiles, ajustement, etc...).

Dans le cas où il est possible de leur adapter un modèle probabiliste, la méthode du maximum de vraisemblance permet de trouver un résultat.

Il nous a paru plus intéressant, à la suite d'un article de MM. Chernoff et Zachs (réf. [2]), d'appliquer, les méthodes bayésiennes à ce type de problème.

Dans une première partie nous rappelons l'essentiel de la statistique bayésienne dans le cadre de la théorie de la décision.

Nous renvoyons le lecteur intéressé par les fondements de celle-ci à un article de M. Morlat (réf. [4]).

La deuxième partie comporte un exposé succinct de la méthode des densités conjuguées (réf. [2] et réf. [5]).

Enfin dans la troisième partie nous appliquons les méthodes précédentes à l'étude d'un éventuel changement de trend dans l'évolution des températures annuelles moyennes relevées au parc St-Maur de 1874 à 1945 (réf. [1]).

-----  
*N.D.L.R. - Ce mémoire rapporte les résultats obtenus par un groupe de travail - composé des cinq signataires, et animé par M. Bernier - dans le cadre du contrôle de connaissance de l'enseignement de l'I.S.U.P., fin de seconde année du cycle supérieur d'Etudes Statistiques, en mai-juin 1969. Il montre, à titre d'exemple, qu'il est possible de faire travailler des étudiants en statistique sur de "vrais" problèmes.*

## I - THEORIE DE LA DECISION BAYESIENNE

### I. 1. a - Introduction

Rappelons qu'un problème de décision comprend :

\* Un espace probabilisable ou espace des états de la nature, noté  $\Theta$  que nous supposons être une partie mesurable de  $R$ .

\* Un espace probabilisable  $\mathcal{X}$  dit espace des observations. En général  $\mathcal{X}$  est tout ou partie de  $R^n$ . Un résultat d'observation ou échantillon est un ensemble de  $n$  observations  $x = (x_1, x_2, \dots, x_n)$ .

A chaque valeur  $\theta$  correspond une probabilité  $P(\cdot | \theta)$  sur  $\mathcal{X}$  : on peut toujours considérer que  $P(\cdot | \theta)$  est une probabilité conditionnelle sachant  $\theta$  sur  $\mathcal{X}$ . On supposera qu'elle admet une densité conditionnelle de  $x$  sachant  $\theta$  notée  $L(x/\theta)$

\* Un espace probabilisable des décisions  $D$ , qui est l'ensemble des décisions permises au statisticien.

\*\* On appelle fonction de décision (ou règle de décision) toute application mesurable de  $\mathcal{X}$  dans  $D$ . Nous noterons  $\delta$  une règle de décision et  $D$  l'ensemble des règles de décision possibles.

Une fonction de perte est une fonction numérique  $W(\theta, d)$  définie sur le produit  $\Theta \times D$ .  $W(\theta, d)$  est la perte subie lorsque l'état de la nature est  $\theta$  et la décision  $d$ .

A toute règle de décision  $\delta$  correspond un risque qui s'écrit :

$$R(\theta, \delta) = \int_{\mathcal{X}} W(\theta, \delta(x)) L(x/\theta) dx.$$

$R(\theta, \delta)$  est la perte moyenne qui résulte de l'emploi de la règle de décision  $\delta$  lorsque l'état de la nature est  $\theta$ .

### I. 1. b - Risque de Bayes

Le statisticien est souvent conduit à traduire ses connaissances a priori sur les états de la nature en introduisant une distribution a priori sur  $\Theta$  appelée distribution à priori des paramètres. On supposera dans cet article que cette distribution admet une densité  $\mu(\theta)$ .

Le propre de la méthode bayésienne est précisément d'utiliser dans un modèle décisionnel une telle distribution a priori choisie indépendamment de l'échantillon observé.

On définit alors le risque de Bayes associé à la densité a priori  $\mu(\theta)$  et à la règle de décision  $\delta$  par :

$$R(\delta) = \int_{\Theta} R(\theta, \delta) \mu(\theta) d\theta = \int_{\Theta \times \mathcal{X}} W(\theta, \delta(x)) L(x/\theta) \mu(\theta) dx d\theta$$

On dit qu'une règle de décision  $\delta_1$  est meilleure qu'une règle de décision  $\delta_2$  pour la densité a priori  $\mu$  si :

$$R(\delta_1) < R(\delta_2)$$

Une règle de décision  $\delta_0$  est optimale au sens de Bayes pour la densité a priori  $\mu(\theta)$  si :

$$R(\delta_0) = \inf_{\delta \in D} R(\delta)$$

Nous rappelons ci-dessous le théorème de Bayes qui permet de transformer utilement l'expression du risque de Bayes.

### I. 1. c - Théorème de Bayes

Si la distribution a priori admet une densité  $\mu(\theta)$  et si  $P(\cdot|\theta)$  admet une densité  $L(x/\theta)$  il existe sur  $\Theta$  une distribution conditionnelle sachant  $x$  de densité :

$$\mu(\theta/x) = \frac{\mu(\theta) L(x/\theta)}{\int_{\Theta} \mu(\theta) L(x/\theta) d\theta}$$

Cette densité est appelée densité a posteriori sur  $\Theta$  sachant  $x$  associée à la densité a priori  $\mu(\theta)$ .

On posera :

$$L \mu(x) = \int_{\Theta} \mu(\theta) L(x/\theta) d\theta$$

Le théorème de Bayes peut alors s'écrire sous la forme suivante dont on fera un usage systématique :

$$L \mu(x) \mu(\theta/x) = L(x/\theta) \mu(\theta) \quad (1)$$

Le risque de Bayes prend alors la forme :

$$R(\delta) = \int_{\Theta \times \mathcal{X}} W(\theta, \delta(x)) L \mu(x) \mu(\theta/x) dx d\theta \quad (2)$$

Les deux paragraphes suivants sont consacrés à la détermination d'une règle de décision optimale dans les problèmes décisionnels particuliers que sont l'estimation et les tests d'hypothèses.

### I. 2 - Estimation Bayésienne

Comme dans la théorie classique de l'estimation, nous supposons que  $\Theta$  est  $R$  ou un borélien de  $R$ . D'autre part  $D = \Theta$  et la fonction de perte s'écrit en posant  $d = \theta'$  :

$$W(\theta, d) = W(\theta, \theta') = (\theta - \theta')^2$$

à un coefficient multiplicatif près.

Comme il est d'usage toute règle de décision  $\delta$  sera appelée estimateur de  $\theta$  et sera notée  $\hat{\theta}$ .

Le risque de Bayes d'un tel estimateur s'écrit compte tenu de (2) :

$$\begin{aligned} R(\hat{\theta}) &= \int_{\Theta \times \mathcal{X}} (\theta - \hat{\theta}(x))^2 L_{\mu(x)} \mu(\theta/x) dx d\theta \\ &= \int_{\mathcal{X}} L_{\mu(x)} dx \int_{\Theta} (\theta - \hat{\theta}(x))^2 \mu(\theta/x) d\theta \end{aligned}$$

La valeur de  $\hat{\theta}(x)$  qui minimise  $\int_{\Theta} (\theta - \hat{\theta}(x))^2 \mu(\theta/x) d\theta$  est :

$$\hat{\theta}(x) = \int_{\Theta} \theta \mu(\theta/x) d\theta.$$

On en déduit que l'estimateur de Bayes pour la densité a priori  $\mu$  est :

$$x \xrightarrow{\hat{\theta}} \int_{\Theta} \theta \mu(\theta/x) d\theta.$$

En d'autres termes  $\hat{\theta}$  est l'espérance mathématique de  $\theta$  pour la mesure a posteriori  $\mu(\theta/x)$ .

### I. 3 - Tests Bayesiens

Classiquement on notera :

$\Theta_0$  : sous-ensemble de  $\Theta$  correspondant à l'hypothèse.

$\Theta_1$  : sous-ensemble de  $\Theta$  correspondant à l'alternative.

$$\Theta = \Theta_0 + \Theta_1$$

$D = \{d_0, d_1\}$  avec  $d_0$  : acceptation de l'hypothèse

$d_1$  : rejet de l'hypothèse

$$\begin{aligned} W(\theta, d_0) &= \begin{cases} 0 & \text{si } \theta \in \Theta_0 \\ c_0 & \text{si } \theta \in \Theta_1 \end{cases} \\ W(\theta, d_1) &= \begin{cases} 0 & \text{si } \theta \in \Theta_1 \\ c_1 & \text{si } \theta \in \Theta_0 \end{cases} \end{aligned}$$

Les règles de décision ou fonctions de tests  $\varphi$  sont des applications à deux valeurs :

$$\varphi(x) = 0 \text{ ou } 1$$

La région d'acceptation  $A$  est l'ensemble des points  $x \in \mathcal{X}$  tels que :

$$\varphi(x) = 0$$

Le risque de Bayes associé à une densité a priori  $\mu(\theta)$  et à une fonction de test  $\varphi$  s'écrit compte tenu de (2) :

$$R(\varphi) = \int_{\mathbf{x}} \varphi(\mathbf{x}) L \mu(\mathbf{x}) d\mathbf{x} \left[ c_1 \int_{\Theta_0} \mu(\theta/\mathbf{x}) d\theta - c_0 \int_{\Theta_1} \mu(\theta/\mathbf{x}) d\theta \right] + c_0 \int_{\Theta_1} \mu(\theta) d\theta$$

Il est clair que le test optimal  $\varphi_0$  est défini par :

$$\varphi_0(\mathbf{x}) = 1 \text{ si } c_1 \int_{\Theta_0} \mu(\theta/\mathbf{x}) d\theta - c_0 \int_{\Theta_1} \mu(\theta/\mathbf{x}) d\theta < 0$$

$$\varphi_0(\mathbf{x}) = 0 \text{ si } c_1 \int_{\Theta_0} \mu(\theta/\mathbf{x}) d\theta - c_0 \int_{\Theta_1} \mu(\theta/\mathbf{x}) d\theta \geq 0$$

Compte tenu de (1) qui permet de transformer les intégrales  $\int_{\Theta_0} \mu(\theta/\mathbf{x}) d\theta$

et  $\int_{\Theta_1} \mu(\theta/\mathbf{x}) d\theta$  il vient :

$$\varphi_0(\mathbf{x}) = 1 \text{ si } \frac{\int_{\Theta_1} L(\mathbf{x}/\theta) \mu(\theta) d\theta}{\int_{\Theta_0} L(\mathbf{x}/\theta) \mu(\theta) d\theta} > \frac{c_1}{c_0}$$

$$\varphi_0(\mathbf{x}) = 0 \text{ ailleurs.}$$

## II - DENSITES CONJUGUEES

### II. 1. a - Généralités

Nous avons vu que l'essentiel de la méthode bayésienne consistait en l'introduction sur  $\theta$  d'une distribution a priori que l'on supposait munie d'une densité  $\mu(\theta)$ .

Cette distribution doit traduire la connaissance qu'a le statisticien des "états de la nature". Il est alors souvent commode de choisir la densité a priori  $\mu(\theta)$  parmi une famille  $F$  de densités satisfaisant aux conditions suivantes :

1 - La famille  $F$  doit se prêter suffisamment au calcul analytique pour permettre le calcul de la densité a posteriori et du risque de Bayes associé.

2 -  $F$  doit être fermée en ce sens que la distribution a posteriori associée à un élément de  $F$  doit encore appartenir à  $F$ .

3 - F doit être assez riche, en d'autres termes doit dépendre de suffisamment de paramètres pour exprimer convenablement les idées a priori du statisticien.

Ces idées doivent en particulier pouvoir se traduire d'une manière simple sur les valeurs des paramètres définissant F.

Ce qui suit a pour but de construire une famille F vérifiant les propriétés ci-dessus pour un problème décisionnel donné.

## II. 1. b - Définition

On appelle conjugée naturelle de la famille  $L(x/\theta)$  indexée par  $\Theta$ , une famille de distributions sur  $\Theta$  indexée par  $\mathfrak{X}$ ,  $L^*(\theta/x)$  définie par :

$$L^*(\theta/x) = L(x/\theta) \cdot g(x)$$

où la fonction réelle  $g(x)$  est une fonction de normalisation vérifiant :

$$\int_{\Theta} g(x) L(x/\theta) d\theta = \int_{\Theta} L^*(\theta/x) d\theta = 1$$

## II. 1. c - Exemples

Eclairons cette définition par quelques exemples.

Exemple 1 - Considérons un échantillon issu d'une loi normale de moyenne  $\theta$  et d'écart-type 1, donc de densité :

$$L(x/\theta) = (2\pi)^{-\frac{n}{2}} \exp - \frac{1}{2} \left[ \sum_{i=1}^n (x_i - \bar{x})^2 + n(\bar{x} - \theta)^2 \right]$$

La conjugée naturelle est une famille de densités sur  $\Theta$  définie par :

$$L^*(\theta/x) = g(x) \exp - \frac{n}{2} (\theta - \bar{x})^2$$

Il s'agit donc d'une famille de lois normales de moyenne  $\bar{x}$  et de variance  $\frac{1}{n}$  ; par suite :

$$g(x) = \left( \frac{n}{2\pi} \right)^{\frac{1}{2}}$$

$\bar{x}$  ne jouant ici aucun rôle autre que celui de paramètre, nous poserons  $\bar{x} = m$  et la conjugée naturelle s'écrit :

$$L^*(\theta/m) = \left( \frac{n}{2\pi} \right)^{\frac{1}{2}} \exp - \frac{n}{2} (\theta - m)^2$$

Exemple 2 - Considérons un n échantillon x issu d'une loi uniforme sur  $[0, \theta]$ .

La densité de cet échantillon s'écrit :

$$L(x/\theta) = \theta^{-n} \delta(\theta - M)$$

avec

$$\delta(a) = \begin{cases} 0 & \text{si } a < 0 \\ 1 & \text{si } a \geq 0 \end{cases} \quad M = \sup_{1 \leq i \leq n} x_i$$

La conjuguée naturelle est de densité

$$L^*(\theta/M) = g(M) \theta^{-n} \delta(\theta - M) \quad \text{avec } M > 0$$

Supposons  $n \geq 2$ . Dans ce cas :

$$\int_{\Theta} L^*(\theta/M) d\theta = 1$$

entraîne :

$$g(M) = (n - 1) M^{n-1}$$

d'où :

$$L^*(\theta/M) = (n - 1) M^{n-1} \theta^{-n} \delta(\theta - M) .$$

## II. 2 - Extension du domaine des paramètres

Les conjuguées naturelles des deux exemples précédents étaient des familles de distribution à un paramètre. Ces familles peuvent être insuffisantes pour décrire correctement les états de la nature.

On appelle conjuguée naturelle enrichie une famille de densités contenant la conjuguée naturelle. En général une densité conjuguée enrichie dépend d'un nombre de paramètres plus grand et la conjuguée naturelle s'obtient en donnant des valeurs particulières à certains paramètres de la conjuguée naturelle enrichie.

Par exemple les densités suivantes :

$$1 - \quad L^*(\theta/m, \alpha) = \left(\frac{\alpha}{2\pi}\right)^{\frac{1}{2}} \exp -\frac{\alpha}{2} (\theta - m)^2$$

$$\alpha > 0 \quad \text{et} \quad -\infty < m < \infty$$

$$2 - \quad L^*(\theta/M, \alpha) = (\alpha - 1) M^{\alpha-1} \theta^{-\alpha} \delta(\theta - M)$$

$$\alpha > \alpha_0 > 1 \quad M > 0$$



constituent respectivement des densités conjuguées naturelles enrichies des exemples 1 et 2.

## II. 3 - Restriction du domaine des paramètres

Au contraire si la conjuguée naturelle fait apparaître un nombre de paramètres trop important pour qu'il soit possible de les estimer tous par notre connaissance des états de la nature, l'on ne considérera qu'une sous famille de la conjuguée naturelle.

## III - APPLICATION A UN PROBLEME CONCRET

### III. 1 - Présentation du modèle

Le présent travail a pour origine une étude de Monsieur J. Vialar sur les températures moyennes annuelles relevées au parc St-Maur pendant la période 1874-1945 (réf. [1]).

L'auteur désirait tester l'hypothèse avancée par de nombreux météorologues d'un réchauffement du climat depuis le début du siècle.

Dans le cas particulier du parc St-Maur et à l'issue d'une étude statistique assez complexe, il concluait à un accroissement des températures de l'ordre de deux degrés par siècle à partir de 1888.

Il nous a paru intéressant d'étudier ce problème sous un tout autre angle en adoptant un modèle simplifié d'évolution des températures. Décrivons le brièvement :

On désigne par  $X_i$  la moyenne des températures relevées au parc St-Maur l'année  $i$  ( $1 \leq i \leq N$ ).

Les variables aléatoires  $X_i$  sont supposées vérifier les conditions suivantes :

\* elles sont mutuellement indépendantes en probabilité

\*  $X_i$  ( $1 \leq i \leq N$ ) est une variable aléatoire normale.

Les deux hypothèses précédentes sont généralement admises lorsqu'on considère des températures moyennes.

Le modèle adopté suppose que l'espérance de  $X_i$  demeure constante pendant les  $T$  premières années,  $T$  étant bien entendu inconnu, et varie ensuite linéairement de  $T$  à  $N$ , l'écart-type restant fixe. Formellement :

$$\begin{aligned} - E(X_i) &= m & V(X_i) &= \sigma^2 & 1 \leq i \leq T \\ - E(X_i) &= m + (i - T)\alpha & V(X_i) &= \sigma^2 & T + 1 \leq i \leq N \end{aligned}$$

Il s'agit de tester l'hypothèse  $T = 1$  (l'espérance des températures varie linéairement depuis 1874) contre l'alternative  $T > 1$  (l'espérance des températures demeure constante entre 1 et  $T$  et varie ensuite linéairement de  $T$  à  $N$ ).

Au cas où le test serait significatif nous nous proposons de déterminer l'époque  $T$  où se produit le changement.

Ce modèle se rapproche de celui traité par Chernoff et Zachs dans les *Annals of Mathematical Statistics* (Réf. [2]). Il en diffère cependant sur les points suivants.

\* Le changement dans les espérances s'effectue à partir de l'époque T d'une manière linéaire donc progressive au lieu de s'effectuer par un saut.

\* Le point de vue adopté est bayésien et fait appel à la méthode générale des densités conjuguées. L'article de Chernoff et Zachs néglige cette technique et utilise simultanément méthodes bayésiennes et méthodes fréquentistes.

### III. 2 - Choix de la densité conjuguée

Sous les hypothèses précédentes la loi du vecteur  $(X_1, X_2, \dots, X_N)$  s'écrit :

$$L(x/T, m, \sigma, \alpha) = (2\pi\sigma^2)^{-N/2} \exp - \frac{1}{2\sigma^2} \left[ \sum_{i=1}^T (x_i - m)^2 + \sum_{i=T+1}^N (x_i - m - (i - T)\alpha)^2 \right]$$

où T, m,  $\alpha$  et  $\sigma$  sont considérés comme des paramètres.

On en déduit l'expression de la densité conjuguée naturelle :

$$L^*(T, m, \alpha, \sigma) = K(2\pi\sigma^2)^{-N/2} \exp - \frac{1}{2\sigma^2} \left[ \alpha^2 \sum_{i=T+1}^N (i - T)^2 + 2m\alpha \sum_{i=T+1}^N (i - T) + Nm^2 - 2mN\bar{x} - 2\alpha \sum_{i=T+1}^N (i - T) x_i + \sum_{i=1}^N x_i^2 \right]$$

Les différentes fonctions de  $x = (x_1, x_2, \dots, x_N)$  figurant dans cette expression doivent être considérées comme des paramètres.

Par suite la densité conjuguée naturelle peut s'écrire sous la forme suivante :

$$L^*(T, m, \alpha, \sigma) = K(2\pi\sigma^2)^{-N/2} \exp - \frac{1}{2\sigma^2} [A(T)\alpha^2 + 2D(T)m\alpha + Nm^2 + 2H(T)\alpha + 2Fm + G]$$

où l'on a posé :

$$A(T) = \frac{(N - T) (N - T + 1) (2N - 2T + 1)}{6}$$

$$D(T) = \frac{(N - T) (N - T + 1)}{2}$$

$$H(T) = - \sum_{i=T+1}^N (i - T)x_i$$

$$F = -N\bar{x} = - \sum_{i=1}^N x_i$$

$$G = \sum_{i=1}^N x_i^2$$

Cette distribution ne se prête guère au calcul du fait que T y figure par l'intermédiaire de fonctions peu maniables A(T), B(T) et H(T).

Nous substituerons donc à  $L^*(T, m, \alpha, \sigma)$  une distribution plus commode  $\mu(T, m, \alpha, \sigma)$  satisfaisant aux hypothèses suivantes :

$$* \mu(T, m, \alpha, \sigma) = \mu_1(T) \times \mu_2(m, \alpha, \sigma)$$

Cette hypothèse exprime l'indépendance des distributions a priori de T et du triplet (m,  $\alpha$ ,  $\sigma$ ). Elle implique que les fonctions A(T), B(T) et H(T) doivent être prises constantes.

On obtient donc :

$$\mu_2(m, \alpha, \sigma) = K \sigma^{-N} \exp - \frac{1}{2\sigma^2} [A\alpha^2 + 2Dm\alpha + Nm^2 + 2H\alpha + 2Fm + G]$$

ou bien en transformant la forme quadratique entre crochets :

$$\begin{aligned} \mu_2(m, \alpha, \sigma) = K \sigma^{-N} \exp - \frac{1}{2\sigma^2} [A(\alpha - \alpha_0)^2 \\ + 2D(m - m_0) (\alpha - \alpha_0) + N(m - m_0)^2 + c] \end{aligned}$$

On enrichit la distribution en introduisant deux nouveaux paramètres  $\nu$  et B. On obtient alors :

$$\begin{aligned} \mu_2(m, \alpha, \sigma) = K \sigma^{-\nu} \exp - \frac{1}{2\sigma^2} [A(\alpha - \alpha_0)^2 \\ + 2D(m - m_0) (\alpha - \alpha_0) + B(m - m_0)^2 + c] \end{aligned}$$

En faisant apparaître dans  $\mu_2(m, \alpha, \sigma)$  la densité marginale de  $\sigma$  et la densité du couple  $(m, \alpha)$  conditionnellement à  $\sigma$  on peut écrire :

$$\mu_2(m, \alpha, \sigma) = K \sigma^{-\nu} \exp - \frac{c}{2\sigma^2} \exp - \frac{1}{2\sigma^2} [A(\alpha - \alpha_0)^2 + B(m - m_0)^2 + 2D(m - m_0)(\alpha - \alpha_0)]$$

Comme nous ignorons la liaison qui existe entre  $\alpha$  et  $m$  conditionnellement à  $\sigma$  nous supposons  $D = 0$ .

Nous prenons pour le paramètre  $T$  une distribution  $\mu_1$  donnant le même poids à l'hypothèse et à l'alternative et traduisant notre incertitude totale sur la date de changement :

$$\mu_1(1) = \frac{1}{2}$$

$$\mu_1(T) = \frac{1}{2} \cdot \frac{1}{N-2} \quad 2 \leq T \leq N-1$$

$$\mu_1(N) = 0$$

En définitive la distribution a priori utilisée s'écrit :

$$\frac{1}{2} K \sigma^{-\nu} \exp - \frac{1}{2\sigma^2} [A(\alpha - \alpha_0)^2 + B(m - m_0)^2 + c] \quad \text{si } T = 1$$

$$\frac{1}{2} \frac{K}{N-2} \sigma^{-\nu} \exp - \frac{1}{2\sigma^2} [A(\alpha - \alpha_0)^2 + B(m - m_0)^2 + c] \quad \text{si } 2 \leq T \leq N-1$$

Les transformations que nous avons fait subir à la densité conjuguée naturelle peuvent paraître excessives.

Notons toutefois qu'il est inutile de conserver une expression trop complexe pour qu'il soit possible d'en estimer les différents paramètres à l'aide des idées a priori que l'on a du phénomène. Notons également que la forme retenue est bien plus riche que la densité que l'on aurait pu construire par produit de distributions indépendantes sur chacun des paramètres.

### III. 3 - Les "a priori" du problème

Avant d'estimer les différents paramètres  $A, B, C, \alpha_0, m_0$  et  $\nu$  de notre distribution a priori, traduisons de manière aussi claire que possible, en insistant au besoin sur leur caractère arbitraire, les idées a priori que nous avons sur ce problème de températures.

(1) Nous estimons que les moyennes annuelles  $X_1$  sont comprises entre 8 et 12 degrés avec la probabilité 0,95.

(2) Nous prenons  $m_0 = 10$  degrés et nous supposons que  $m$ , exprimé en degrés, est compris entre 9,5 et 10,5 avec la probabilité 0,95.

(3) L'écart-type  $\sigma$  a un mode égal à 1 et une probabilité de 0,95 d'être inférieur à 2.

(4) Le paramètre  $\alpha$  a une très forte probabilité d'être positif. Nous considérons que la probabilité qu'il soit négatif est de l'ordre de 0,01.

La moyenne  $\alpha_0$  de  $\alpha$  est de l'ordre de 0,03 degrés ce qui correspond à une variation de 3 degrés par siècle.

Ce dernier chiffre est le plus contestable ; aussi avons-nous également essayé les valeurs 0,02 et 0,04.

### III. 4 - Estimation des paramètres de la densité conjuguée

Les idées précédentes concernent essentiellement les distributions marginales des paramètres  $m$ ,  $\alpha$  et  $\sigma$ .

Un calcul classique fournit les résultats suivants :

\*  $\frac{c}{2\sigma^2}$  suit une loi  $\Gamma$  à  $\frac{\nu - 3}{2}$  degrés de liberté.

La loi de  $\sigma$  est donc une distribution gamma inverse de type 2 de paramètres  $\nu' = \nu - 3$  et  $\delta' = \sqrt{\frac{c}{\nu - 3}}$  (cf. référence [3]).

Le mode d'une telle distribution est  $\delta' \sqrt{\frac{\nu'}{\nu' + 1}} = \sqrt{\frac{c}{\nu - 2}}$ .

\*  $\sqrt{\nu - 3} \cdot \sqrt{\frac{B}{c}} (m - m_0)$  et  $\sqrt{\nu - 3} \sqrt{\frac{A}{c}} (\alpha - \alpha_0)$  suivent des lois de Student à  $\nu - 3$  degrés de liberté.

La première partie de la condition (3) fournit la relation  $c = \nu - 2$ .

Pour traduire la deuxième partie nous essayons les différentes valeurs entières possibles de  $\nu$ , c'est-à-dire les entiers strictement supérieurs à 3. On retiendra évidemment la première valeur de  $\nu$  pour laquelle :

$$P[\sigma \in (0, 2)] \geq 0,95$$

On a utilisé les tables de la loi du  $\chi^2$  puisque  $\frac{\nu - 2}{\sigma^2}$  suit une loi du  $\chi^2$  à  $\nu - 3$  degrés de liberté.

On trouve  $\nu = 10$  donc  $c = 8$ .

\* Pour traduire la condition (2), il est nécessaire de construire pour le paramètre  $m$  un intervalle de probabilité à 95 % centré sur  $m_0$ . On trouve facilement qu'un tel intervalle s'écrit :

$$\left[ m_0 - 2,365 \sqrt{\frac{c}{B(\nu - 3)}}, m_0 + 2,365 \sqrt{\frac{c}{B(\nu - 3)}} \right]$$

On en déduit la condition :

$$0,5 = 2,365 \sqrt{\frac{8}{7B}}$$

c'est-à-dire

$$B = 25,57$$

\* La première partie de la condition (4) s'écrira :

$$P(\alpha < 0) = P(\sqrt{\nu - 3} \sqrt{\frac{A}{c}} (\alpha - \alpha_0) < -\alpha_0 \sqrt{\frac{A}{c}} \sqrt{\nu - 3}) = 0,01$$

d'où

$$\alpha_0 \sqrt{\nu - 3} \sqrt{\frac{A}{c}} = 2,998$$

ou encore

$$A \alpha_0^2 \simeq 10$$

### III. 5 - Test de l'hypothèse $T = 1$ contre l'alternative $T > 1$

Rappelons que l'espace des paramètres s'écrit  $\Theta = \mathbb{R}^2 \times \mathbb{R}^+ \times \{1, 2, \dots, N - 1\}$  le paramètre  $\theta$  étant le quadruplet  $(m, \alpha, \theta, T)$  dans lequel  $m$  et  $\alpha$  appartiennent à  $\mathbb{R}$ ,  $\sigma$  à  $\mathbb{R}^+$  et  $T$  à l'ensemble  $\{1, 2, \dots, N - 1\}$ .

Pour alléger les notations on notera  $\xi$  le triplet  $(m, \alpha, \sigma)$  et  $\theta$  sera noté  $(\xi, T)$ .

L'hypothèse s'écrit :  $\Theta_0 = \mathbb{R}^2 \times \mathbb{R}^+ \times \{1\}$

L'alternative :  $\Theta_1 = \mathbb{R}^2 \times \mathbb{R}^+ \times \{2, \dots, N - 1\}$

Sous l'hypothèse  $\Theta_0$ , la vraisemblance de l'échantillon s'écrit :

$$L(x/\theta_0) = (2\pi \sigma^2)^{-N/2} \exp - \frac{1}{2\sigma^2} \left[ \sum_{i=1}^N (x_i - m)^2 \right] \text{ avec } \theta_0 = (\xi, 1)$$

Sous l'alternative  $\Theta_1$ , la vraisemblance de l'échantillon conditionnellement à  $\theta_1 = (\xi, T)$  ( $2 \leq T \leq N - 1$ ) s'écrit :

$$L(x/\theta_1) = (2\pi \sigma^2)^{-N/2} \exp - \frac{1}{2\sigma^2} \left[ \sum_{i=1}^N (x_i - m)^2 + \sum_{i=T+1}^N (x_i - m - (i - T)\alpha)^2 \right]$$

Le test nécessite le calcul de  $\int_{\Theta_1} L(x/\theta) \mu(\theta) d\theta$  et de  $\int_{\Theta_0} L(x/\theta) \mu(\theta) d\theta$  (cf. I.3) dans lequel

$$\mu(\theta) = K \sigma^{-\nu} \left[ \exp - \frac{1}{2\sigma^2} (A(\alpha - \alpha_0)^2 + B(m - m_0)^2 + c) \right] \times \mu_1(T)$$

$$\mu_1(T) = \begin{cases} \frac{1}{2} \text{ si } T = 1 & \text{(hypothèse)} \\ \frac{1}{2(N-2)} \text{ si } 2 \leq T \leq N-1 & \text{(alternative)} \end{cases}$$

On trouve sans difficultés :

$$\int_{\Theta_1} L(x/\theta) \mu(\theta) d\theta = \frac{H}{2(N-2)} \sum_{T=2}^{N-1} \frac{1}{(A+U_T)^{\frac{1}{2}}} \left[ k - \frac{g_T^2}{A+U_T} + c \right]^{-\frac{N+v-3}{2}}$$

$$\int_{\Theta_0} L(x/\theta) \mu(\theta) d\theta = \frac{H}{2} \frac{1}{(A+U_1)^{\frac{1}{2}}} \left[ k - \frac{g_1^2}{A+U_1} + c \right]^{-\frac{N+v-3}{2}}$$

dans lequel :

$$U_T = \frac{(N-T)(N-T+1)(2N-2T+1)}{6} - \frac{(N-T)^2(N-T+1)^2}{4(B+N)}$$

$$g_T = A \alpha_0 + \sum_{i=T+1}^N (i-T) x_i - \frac{(N-T)(N-T+1)}{2(B+N)} [N\bar{x} + Bm_0]$$

$$k = \sum_{i=1}^N x_i^2 + B m_0^2 - \left( \frac{N\bar{x} + B m_0}{(B+N)^{\frac{1}{2}}} \right)^2 + A \alpha_0^2$$

H : constante de normalisation.

La région critique du test est définie par :

$$\frac{1}{N-2} \sum_{T=2}^{N-1} \left( \frac{A+U_1}{A+U_T} \right)^{\frac{1}{2}} \left[ \frac{k - \frac{g_1^2}{A+U_1} + c}{k - \frac{g_T^2}{A+U_T} + c} \right]^{\frac{N+v-3}{2}} > \frac{c_0}{c_1}$$

En prenant la même fonction de perte sous l'hypothèse et l'alternative ( $c_0 = c_1$ ), on obtient :

$$\frac{1}{N-2} \sum_{T=2}^{N-1} \left( \frac{A+U_1}{A+U_T} \right)^{\frac{1}{2}} \left[ \frac{k - \frac{g_1^2}{A+U_1} + c}{k - \frac{g_T^2}{A+U_T} + c} \right]^{\frac{N+v-3}{2}} > 1$$

### III. 6 - Estimation bayésienne de T

Il se pose maintenant le problème de l'estimation de T. Les paramètres  $\xi$  et T étant indépendants, la vraisemblance de l'échantillon conditionnellement à T s'écrit :

$$L(x/T) = \int_{R^2 \times R^+} L(x/(\xi, T)) \mu_2(\xi) d\xi$$

où

$$\mu_2(\xi) = \mu_2(m, \alpha, \sigma) = K \sigma^{-\nu} \exp - \frac{1}{2\sigma^2} \left[ A(\alpha - \alpha_0)^2 + B(m - m_0)^2 + c \right]$$

Le calcul est identique à ceux effectués au paragraphe précédent à cela près que l'on n'intègre pas par rapport à  $\mu_1$

On trouve :

$$L(x/T) = \frac{H}{(A + U_T)^{\frac{1}{2}}} \left[ k - \frac{g_T^2}{A + U_T} + c \right]^{-\frac{N+\nu-3}{2}}$$

La loi a posteriori de T s'écrit d'après le théorème de Bayes :

$$L(T/x) = \frac{\frac{1}{N-2} L(x/T)}{L(x/1) + \frac{1}{N-2} \sum_{T=2}^{N-1} L(x/T)} \quad 2 \leq T \leq N-1$$

$$L(1/x) = \frac{L(x/1)}{L(x/1) + \frac{1}{N-2} \sum_{T=2}^{N-1} L(x/T)} \quad T = 1$$

L'estimateur  $\hat{T}(x)$  de T est l'espérance mathématique de T pour la loi a posteriori de T. Il s'écrit :

$$\hat{T}(x) = \frac{L(x/1) + \frac{1}{N-2} \sum_{T=2}^{N-1} T L(x/T)}{L(x/1) + \frac{1}{N-2} \sum_{T=2}^{N-1} L(x/T)}$$

#### IV - CONCLUSIONS

##### IV. 1 - Présentation des résultats

Les calculs précédents ont été réalisés avec le concours du groupe Statistiques et Calculs du Centre d'Etudes Nucléaires de Fontenay-aux-Roses ; les valeurs des différents paramètres étant celles définies au paragraphe III.4

Pour tester la sensibilité du modèle par rapport à  $\alpha_0$ , nous avons introduit les valeurs 0,02, 0,03 et 0,04.

Les résultats sont les suivants, en appelant TEST la statistique :

$$TEST = \frac{1}{N-2} \sum_{T=2}^{N-2} \frac{A + U_1}{A + U_T} \left( \frac{k - \frac{g_1^2}{A + U_1} + c}{k - \frac{g_T^2}{A + U_T} + c} \right)^{\frac{N+\nu-3}{2}}$$



$$1/ \alpha_0 = 0,02$$

TEST = 2,36 (test significatif)

$\hat{T} = 13$  correspondant à un changement en 1886.

$$2/ \alpha_0 = 0,03$$

TEST = 11,21 (test significatif)

$\hat{T} = 24$  correspondant à un changement en 1897.

$$3/ \alpha_0 = 0,04$$

TEST = 26,48 (test significatif)

$\hat{T} = 31$  correspondant à un changement en 1904.

On remarque une sensibilité certaine du modèle par rapport à  $\alpha_0$ , espérance de la loi a priori sur  $\alpha$ .

Notre information insuffisante sur  $\alpha_0$  explique la plage importante que nous obtenons sur les estimations  $\hat{T}$  de  $T$ .

La valeur trouvée pour  $\alpha_0 = 0,02$  est très voisine de celle estimée par Monsieur J. Vialar. Toutefois la sensibilité du modèle ne nous permet guère d'être très affirmatifs sur cette date.

Tout au plus peut-on dire, et la lecture du graphique de température en annexe ne nous contredira pas, qu'il y a un changement dans l'évolution des températures à la fin du XIX<sup>e</sup> siècle.

#### IV. 2 - Conclusion

Le travail présenté n'est nullement exhaustif. Il eût été logique en particulier de calculer la distribution a posteriori de  $\alpha$  et d'en déduire son estimation bayésienne ce qui est en fait le problème symétrique de celui traité pour  $T$ .

Analytiquement, les calculs aboutissaient à des intégrales irréductibles nécessitant un calcul approché sur ordinateur.

D'autre part il est clair que les lois a posteriori de  $\alpha$  et  $T$  sont fortement liées. Les variations de l'estimation a posteriori  $\hat{T}$  se répercuteront donc sur l'estimation a posteriori  $\hat{\alpha}$ .

Nous avons jugé qu'une telle étude ne pouvait qu'augmenter la part déjà importante réservée aux calculs analytiques sans rien apporter d'essentiel à la compréhension de la méthode.

Nous nous sommes efforcés en effet de présenter au lecteur une application de la méthode bayésienne plutôt qu'une analyse complète d'un problème particulier. Dans les modèles chronologiques où se manifeste un brusque changement de la valeur d'un paramètre à une date inconnue, les méthodes bayésiennes semblent bien adaptées.

Il ne faut pas se dissimuler cependant que la démarche bayésienne suppose le choix d'une distribution a priori bien précise. L'article montre bien à quelle dispersion des résultats l'on peut s'attendre faute d'une information a priori suffisante.

## BIBLIOGRAPHIE

- [1] Jean VIALAR : Etude statistique des températures de Paris St-Maur pour la période 1874-1945.  
Journal Scientifique de la Météorologie IV - 13.
- [2] CHERNOFF et ZACHS: Estimating the current mean of a normal distribution which is subject to changes in time.  
Annals of Mathematical Statistics 1964, pages 1013-1018.
- [3] RAIFFA et SCHLAIFER: Applied statistical decision theory  
Harvard University 1961.
- [4] MORLAT : Statistique et théorie de la décision  
Revue de Statistique Appliquée (1964, 2ème tr.).
- [5] HANSEL et GROUCHKO : Prévission séquentielle par la méthode de Bayes  
Revue de Statistique Appliquée (1965, 3ème tr.).

## ANNEXE

Températures moyennes annuelles relevées au  
parc St Maur de 1874 à 1945

Année	Température en 1/100 de degré	Année	Température en 1/100 de degré	Année	Température en 1/100 de degré	Année	Température en 1/100 de degré
1874	1029	1892	1015	1910	1034	1928	1138
1875	1019	1893	1084	1911	1125	1929	1052
1876	1059	1894	1043	1912	1048	1930	1122
1877	1045	1895	985	1913	1086	1931	1031
1878	1004	1896	984	1914	1054	1932	1085
1879	816	1897	1057	1915	1036	1933	1066
1880	1058	1898	1066	1916	1040	1934	1172
1881	981	1899	1081	1917	933	1935	1099
1882	1014	1900	1112	1918	1048	1936	1089
1883	998	1901	1001	1919	985	1937	1136
1884	1055	1902	987	1920	1059	1938	1119
1885	976	1903	1032	1921	1150	1939	1092
1886	1031	1904	1040	1922	1001	1940	1025
1887	881	1905	988	1923	1072	1941	1041
1888	895	1906	1053	1924	1028	1942	1044
1889	947	1907	1013	1925	1019	1943	1154
1890	932	1908	977	1926	1100	1944	1084
1891	946	1909	947	1927	1051	1945	1180

