

P. BETHOUX

P. CHEMARIN

Quelques problèmes linéaires de statistique concernant un processus stochastique de moyenne inconnue

Publications du Département de Mathématiques de Lyon, 1967, tome 4, fascicule 2, p. 149-166

http://www.numdam.org/item?id=PDML_1967__4_2_A5_0

© Université de Lyon, 1967, tous droits réservés.

L'accès aux archives de la série « Publications du Département de mathématiques de Lyon » 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/>

QUELQUES PROBLEMES LINEAIRES DE STATISTIQUE CONCERNANT UN PROCESSUS STOCHASTIQUE DE MOYENNE INCONNUE

Soit $\{x_t, t \in T\}$ un processus stochastique du second ordre pour une famille de probabilités, de manière que :

(i) on peut prendre la moyenne ρ du processus comme paramètre repérant un élément de la famille de probabilités considérée.

(ii) la covariance de $\{x_t, t \in T\}$ sous la loi P_ρ

$$E_\rho [(x_t - \rho(t))(x_s - \rho(s))]$$

est une fonction numérique sur $T \times T$, $R(t,s)$, indépendante de ρ et connue.

Plusieurs auteurs ont étudié les divers problèmes de Statistique qui peuvent se poser à partir de l'observation d'une réalisation du processus stochastique considéré, lorsque l'on suppose que ρ appartient à priori à un sous-espace vectoriel donné, de dimension finie de l'espace des fonctions numériques définies sur T . Voir par exemple [1] et [2].

Nous avons essayé ici de reprendre ces mêmes problèmes en supposant que ρ appartient à une partie quelconque de l'espace auto-reproduisant associé à $R(t,s)$.

Soient

- $\{x_t, t \in T\}$ un processus stochastique réel ou complexe sur $(\Omega, \mathcal{F})$.

- M un ensemble de fonctions numériques définies sur T .

- Une famille de probabilités sur $(\Omega, \mathcal{F})$, (P_ρ) $\rho \in M$ telle que, si $E_\rho(x)$ désigne l'espérance mathématique par rapport à P_ρ de la variable aléatoire x définie sur $(\Omega, \mathcal{F})$, on ait :

$$- E_\rho(x_t) = \rho(t) \quad \text{pour tout } t \in T \text{ et pour tout } \rho \in M$$

$$- E_\rho[(x_t - \rho(t))(x_s - \rho(s))] = R(t, s) \quad \text{pour tous les } s \text{ et } t \in T \text{ et pour tout } \rho \in M.$$

Nous envisageons dans la suite les problèmes suivants : la fonction $R(t, s)$ et l'ensemble M sont connus. En supposant que le processus $\{x_t, t \in T\}$ est régi par l'une (inconnue) des lois P_ρ , on cherche au vu d'une réalisation de $\{x_t, t \in T\}$ à estimer $\rho(t_0)$ pour $t_0 \in T$, un paramètre numérique plus général fonction de ρ , ou encore la valeur prise par une variable aléatoire y dont, sous la loi P_ρ , la moyenne est une fonctionnelle connue de ρ et dont la variance et la covariance avec x_t sont indépendantes de ρ et données.

I - QUELQUES DEFINITIONS ET LEMMES

Espace auto-reproduisant lié à R

Pour $s \in T$, désignons par R^s la fonction définie sur T par :

$$R^s(t) = R(t, s) \quad (I.1)$$

L'espace vectoriel engendré par les R^s avec $s \in T$, est un espace préhilbertien $\mathcal{E}$ avec le produit scalaire défini par :

$$\langle R^s, R^t \rangle = R(t, s) \quad (I.2)$$

La convergence en norme dans $\mathcal{E}$ entraîne la convergence ponctuelle, donc l'espace de Hilbert $\Phi(R)$ complété de $\mathcal{E}$ est constitué par l'ensemble des fonctions numériques sur T qui sont des limites ponctuelles de suites de Cauchy en norme d'éléments de $\mathcal{E}$. $\Phi(R)$ est appelé l'espace auto-reproduisant associé à R . Voir par exemple [2] et [4].

De plus, on peut constater :

(i) que pour chaque $\varphi \in \Phi(R)$, on a

$$\varphi(t) = \langle \varphi, R^t \rangle \text{ pour tout } t \in T \quad (I.3)$$

(ii) qu'une condition nécessaire et suffisante pour qu'une fonction numérique φ définie sur T appartienne à $\Phi(R)$ est qu'il existe une constante K_φ telle que

$$\left| \sum c_\nu \varphi(t_\nu) \right|^2 \leq K_\varphi \sum_{\nu, \nu'} c_\nu \overline{c_{\nu'}} R(t_\nu, t_{\nu'}) \quad (I.4)$$

quels que soient la famille finie $\{t_\nu\}$ de valeurs de T et les nombres c_ν .

Définitions I.1 : "On désignera

- " -par X l'espace vectoriel engendré par les x_t pour $t \in T$
- " -par Z le quotient, par la relation d'équivalence "égale
- " presque-sûrement P_ρ pour tout $\rho \in M$, de l'espace des
- " variables aléatoires définies sur $(\Omega, \mathcal{F})$ qui sont limites,
- " en moyenne quadratique P_ρ pour tout $\rho \in M$, d'une suite
- " d'éléments de X .

Lemme I.1 : "Si $M \subset \Phi(R)$

$$(x, y) = E_\rho [(x - E_\rho(x))(y - E_\rho(y))] \quad (I.5)$$

- " définit sur Z un produit scalaire avec lequel Z est un
- " espace de Hilbert.

- " $\rho \rightarrow E_\rho(x)$ est une forme linéaire bornée sur Z .

Notons que pour tous les x et $y \in Z$

$$(I.5) \quad E_{\rho} [(x - E_{\rho}(x))(y - E_{\rho}(y))] \text{ ne dépend pas de } \rho \in M.$$

De (I.4) on déduit que pour tout $x \in X$,

$$|E_{\rho}(x)|^2 \leq K_{\rho}(x, x) \quad (I.6)$$

Cette relation se prolonge à tout $z \in Z$ (I.6'), et alors

$$E_{\rho}(|z|^2) = \text{var}(z) + |E_{\rho}(z)|^2 \leq (1 + K_{\rho})(z, z) \quad (I.7)$$

ce qui montre que

$(x, y) = \text{cov}(x, y)$ est une forme hermitienne, définie positive sur Z .

Etablissons que Z muni du produit scalaire (I.5) est complet.

Soit $\{x_n\}$, $n = 1, 2, \dots$ une suite de Cauchy au sens de la norme $\|\cdot\|$ dérivant de ce produit scalaire.

Il résulte de (I.7) que x_n est de Cauchy en moyenne quadratique P_{ρ} , pour tout $\rho \in M$.

Soit alors $\hat{x}_n$ un représentant de x_n . D'après l'inégalité de Tchébichev, pour tout $\xi > 0$

$$P_{\rho} \left\{ \omega : |\hat{x}_n(\omega) - \hat{x}_m(\omega)| \geq \xi \right\} \leq \frac{1}{\xi^2} E_{\rho} |\hat{x}_n - \hat{x}_m|^2 \quad (I.8)$$

$$\leq \frac{1 + K_{\rho}}{\xi^2} \|x_n - x_m\|^2 \text{ compte tenu de (I.7)} \quad (I.9)$$

Alors suivant un procédé classique, soit n_k une suite strictement croissante d'entiers telle que

$$\|x_{n_{k+1}} - x_{n_k}\|^2 < \frac{1}{2^{3k}} \quad \text{pour tout } k \in \mathbb{N} \quad (I.10)$$

Si

$$M_k = \left\{ \omega : |\hat{x}_{n_{k+1}}(\omega) - \hat{x}_{n_k}(\omega)| > \frac{1}{2^k} \right\} \quad (I.11)$$

on constate que $\hat{x}_{n_k}(\omega)$ converge sur $B = \liminf_{k \rightarrow +\infty} M_k^c$ et que

$$P_{\rho}(B^c) = 0 \quad \text{pour tout } \rho \in M.$$

Si on pose

$$\begin{aligned} \mathfrak{X}(\omega) &= \lim_{k \rightarrow +\infty} \mathfrak{X}_{n_k}(\omega) & \text{si } \omega \in B \\ \mathfrak{X}(\omega) &= 0 & \text{si } \omega \notin B \end{aligned} \quad (\text{I.12})$$

$\mathfrak{X}_{n_k} \rightarrow \mathfrak{X}$ presque-sûrement P_ρ pour tout $\rho \in M$.

Il en résulte que $\mathfrak{X}_n \rightarrow \mathfrak{X}$ en moyenne quadratique P_ρ pour tout $\rho \in M$ et comme chaque $\mathfrak{X}_n$ est limite en moyenne quadratique P_ρ pour tout $\rho \in M$ d'éléments de X , il en est de même de $\mathfrak{X}$.

$\mathfrak{X}$ est donc un représentant d'un élément x de Z et

$$\|x - x_n\|^2 = \text{var}(\mathfrak{X} - \mathfrak{X}_n) \rightarrow 0 \quad \text{quand } n \rightarrow +\infty.$$

La relation (I.6') exprime que la forme linéaire sur Z

$$x \rightarrow E_\rho(x) \quad \text{est bornée.}$$

Lemme I.2 : "L'application I qui à $z \in Z$ associe la fonction

" numérique définie sur T par

$$" \quad (Iz)(t) = (x_t, z) \quad (\text{I.13})$$

" est une bijection de Z sur $\Phi(R)$ et

$$" \quad \langle Iz_1, Iz_2 \rangle = (z_2, z_1) \quad (\text{I.14})$$

En remarquant que

$$Ix_s = R^s \quad \text{pour tout } s \in T \quad (\text{I.15})$$

$$\langle R^s, R^t \rangle = (x_t, x_s) \quad \text{pour tous les } s \text{ et } t \in T \quad (\text{I.16})$$

$$I\left(\sum_i \alpha_i x_{s_i}\right) = \sum_i \overline{\alpha}_i Ix_{s_i} \quad (\text{I.17})$$

on déduit immédiatement que I est une bijection du quotient X^* de X par la relation d'équivalence "égale presque-sûrement P_ρ pour tout $\rho \in M$ " sur $\mathcal{C}$ et que (I.14) est vraie pour z_1 et z_2 appartenant à X^* .

Soit z quelconque $\in Z$, $z = \lim_{n \rightarrow +\infty} z_n$ avec $z_n \in X^*$.

On constate immédiatement que Iz est limite ponctuelle des fonctions Iz_n constituant une suite de Cauchy en norme. Donc Iz

est limite en norme de Iz_n , par suite (I.14) est vraie pour z_1 et z_2 quelconques appartenant à Z et I est une bijection de Z sur $\Phi(R)$.

Lemme 1.3 : "Soient :

" - V le sous-espace de Hilbert de Z engendré par les

" $v_\rho = I^{-1}(\rho)$, $\rho \in M$.

" - x' la projection orthogonale dans Z de $x \in Z$ sur V .

" On a :

" $\left\{ \begin{array}{l} E \rho(x') = E \rho(x) \end{array} \right.$ pour tout $\rho \in M$.

" (I.19) $\left\{ \begin{array}{l} \text{var}(x') \leq \text{var}(x) \end{array} \right.$ l'égalité n'ayant lieu que si $x \in V$.

x' est défini par

$x' \in V$ et $([x - x'], v_\rho) = 0$ pour tout $\rho \in M$.

D'où $E \rho(x') = E \rho(x)$ pour tout $\rho \in M$.

Le reste se déduit de l'application du théorème de Pythagore :

$$\text{var}(x) = \|x\|^2 = \|x'\|^2 + \|x - x'\|^2 = \text{var } x' + \|x - x'\|^2$$

II - ESTIMATION LINEAIRE

Définition II.1 : "Les éléments de Z seront appelés les estimateurs linéaires à partir de l'observation de $\{x_t, t \in T\}$."

A - Estimation linéaire d'un paramètre numérique

Soit $\alpha = h(\rho)$ un paramètre numérique à estimer.

Théorème II.1 : "Si, pour comparer les estimateurs linéaires, on

" prend un critère tel que :

" (II.1) $\left\{ \begin{array}{l} E \rho(x') = E \rho(x) \\ \text{var}(x') < \text{var}(x) \end{array} \right.$ pour tout $\rho \in M$

" entraîne x' meilleur que x , alors on peut se limiter à la

" famille des estimateurs linéaires qui appartiennent à V .

Ceci est la traduction du lemme I.3.

Remarque : Il en est en particulier ainsi, si l'on cherche à minimiser la fonction $\rho \rightarrow E_{\rho} |x - h(\rho)|^2$.

$$\text{En effet, } E_{\rho} |x - h(\rho)|^2 = \text{var } x + |E_{\rho}(x) - h(\rho)|^2 \quad (\text{II.2})$$

Théorème II.2 : "Une condition nécessaire et suffisante pour que

- " $\alpha = h(\rho)$ admette un estimateur linéaire non biaisé est que
- " h soit la restriction à M d'une forme linéaire et bornée sur
- " le sous-espace V' de $\Phi(R)$ engendré par M .
- " La différence de deux estimateurs linéaires non biaisés est
- " orthogonale à V .

Supposons que $\alpha = h(\rho)$ admette un estimateur linéaire $\hat{\alpha}$ tel que

$$\begin{aligned} E_{\rho}(\hat{\alpha}) &= h(\rho) \quad \text{pour tout } \rho \in M \\ \text{alors } h(\rho) &= (\hat{\alpha}, v_{\rho}) \\ &= \langle \rho, I \hat{\alpha} \rangle \quad \text{pour tout } \rho \in M \end{aligned} \quad (\text{II.3})$$

ce qui exprime que $h(\rho)$ est la restriction à M de la forme linéaire et continue sur $\Phi(R)$

$$\varphi \rightarrow \langle \varphi, I \hat{\alpha} \rangle.$$

Réciproquement, supposons que $h(\cdot)$ soit la restriction à M d'une forme linéaire et continue sur V' . Alors il existe $g \in V'$ tel que

$$\begin{aligned} h(\rho) &= \langle \rho, g \rangle \quad \text{pour tout } \rho \in M \\ &= (I^{-1}g, v_{\rho}) = E_{\rho}(I^{-1}g) \end{aligned} \quad (\text{II.4})$$

$I^{-1}g$ est un estimateur linéaire non biaisé (appartenant d'ailleurs à V) de $\alpha = h(\rho)$.

Si $\hat{\alpha}_1$ et $\hat{\alpha}_2$ sont deux estimateurs linéaires non biaisés de $\alpha = h(\rho)$,

$$(\hat{\alpha}_1, v_{\rho}) = E_{\rho}(\hat{\alpha}_1) = E_{\rho}(\hat{\alpha}_2) = (\hat{\alpha}_2, v_{\rho}) \quad \text{pour tout } \rho \in M$$

D'où $\hat{\alpha}_1 - \hat{\alpha}_2$ est orthogonal à V .

Remarque : D'après la dernière assertion du théorème II.2 et le lemme I.3, s'il existe un estimateur linéaire non biaisé de $\alpha = h(\rho)$, il y en a un seul de variance minimum.

Cas particuliers :

1) Estimation de $\rho(t_0)$ avec $(t_0 \in T)$

$\rho \rightarrow \rho(t_0)$ est la restriction à M de la forme linéaire et continue sur $\Phi(R)$:

$$\varphi \rightarrow \langle \varphi, R^{t_0} \rangle$$

On aurait d'ailleurs pu exhiber un estimateur linéaire non biaisé évident : x_{t_0} .

L'estimateur linéaire non biaisé de variance minimum est la projection orthogonale dans Z de x_{t_0} sur V .

2) Si M est un sous-espace vectoriel de $\Phi(R)$, de dimension finie, alors toute forme linéaire sur M admet des estimateurs linéaires non biaisés et parmi ceux-ci un seul de variance minimum. Ce cas a été étudié en détail. Voir [1].

3) Supposons que M soit un sous-espace de Hilbert séparable de $\Phi(R)$ (Remarquons que dans la plupart des cas pratiques $\Phi(R)$ est séparable).

Désignons par (ρ_j) $j = 1, 2, \dots$ une base de $M = V'$.
Ecrivons :

$$\rho = \sum_{j=1}^{+\infty} a_j \rho_j \quad (\text{II.5})$$

$a_j(\rho)$ est une forme linéaire et continue sur M . Elle admet des estimateurs linéaires non biaisés. Désignons par $\hat{a}_j$ celui de variance minimum ($\hat{a}_j \in V$).

Soit (u_j) la base conjuguée de la base (ρ_j)
 $u_j \in M$, $\langle \rho_i, u_j \rangle = \delta_{i,j}$ $i, j = 1, 2, \dots$

On a :

$$a_j(\rho) = \langle \rho, u_j \rangle$$

D'où $\hat{a}_j = I^{-1}(u_j)$ (II.6)

On a alors le théorème suivant :

Théorème II.3 : "L'estimateur linéaire et non biaisé de variance

" minimum de toute forme $h(\rho)$ linéaire et continue sur M ,

" est donné par :

"
$$\hat{h} = \sum_{j=1}^{+\infty} \hat{a}_j h(\rho_j)$$

$h(\rho)$ étant une forme linéaire et continue sur M , soit H l'unique élément de M tel que

$$h(\rho) = \langle \rho, H \rangle \quad \text{pour tout } \rho \in M \quad (\text{II.7})$$

L'estimateur linéaire et non biaisé, de variance minimum de $h(\rho)$ est

$$I^{-1}(H)$$

H s'écrit sur la base u_j de M

$$H = \sum_{j=1}^{+\infty} \langle H, \rho_j \rangle u_j = \sum_{j=1}^{+\infty} \overline{h(\rho_j)} u_j$$

Compte tenu de la continuité de I^{-1} et de son anti-linéarité

$$\begin{aligned} I^{-1}(H) &= \sum_{j=1}^{+\infty} h(\rho_j) I^{-1}(u_j) \\ &= \sum_{j=1}^{+\infty} \hat{a}_j h(\rho_j) \end{aligned} \quad (\text{II.8})$$

Remarque : $\hat{h}_n = \sum_{j=1}^{+n} \hat{a}_j h(\rho_j)$ tend quand $n \rightarrow +\infty$ vers $\hat{h}$ dans Z , ce qui implique d'après (I.7) que $\hat{h}_n \rightarrow \hat{h}$ en moyenne quadratique P_ρ pour tout $\rho \in M$. Il sera donc légitime d'approcher $\hat{h}$ par $\hat{h}_n$ qui est l'estimateur linéaire et non biaisé, de variance minimum de la forme h_n égale à h sur le sous-espace M_n engendré par $\rho_1, \dots, \rho_n$ et nulle sur $M \ominus M_n$.

Exemple : Considérons l'exemple suivant :

- T est l'intervalle $[0, T]$, x_t est réel.
- $R(t, s) = \min(t, s)$ pour $0 \leq t, s \leq T$

On sait ([1]) que $\Phi(R)$ est constitué des fonctions absolument continues sur $[0, T]$ et de dérivée de carré intégrable sur $[0, T]$ par rapport à la mesure de Lebesgue et que

$$\langle f, g \rangle = \int_0^T f'(t)g'(t)dt \quad \text{pour } f \text{ et } g \in \Phi(R)$$

$\Phi(R)$ est séparable et une base orthogonale est constituée par les éléments φ_n suivants :

$$\begin{cases} \varphi_0(t) = t \\ \varphi_{2k}(t) = \int_0^t \cos\left(\frac{2\pi k}{T} u\right) du \\ \varphi_{2k-1}(t) = \int_0^t \sin\left(\frac{2\pi k}{T} u\right) du \end{cases} \quad k \geq 1$$

Supposons que M soit constitué des fonctions de $\Phi(R)$ périodiques de période $\frac{T}{m}$ (m entier fixé). Ce problème a été considéré par H.B. Mann dans [3].

M est un sous-espace vectoriel de $\Phi(R)$. Il est fermé car la convergence au sens de la norme de $\Phi(R)$ entraîne la convergence ponctuelle et toute limite ponctuelle de fonctions de période $\frac{T}{m}$ est de période $\frac{T}{m}$.

Or une base de M est constituée par les fonctions,

$$\varphi_{2m}, \varphi_{2m-1}, \varphi_{4m}, \varphi_{4m-1}, \dots$$

D'autre part

$$I^{-1}(f) = \int_0^T f'(u) dx_u$$

$\rho \in M$ s'écrit :

$$\rho = \sum_{k=1}^{+\infty} (a_k \varphi_{2km} + b_k \varphi_{2km-1})$$

$$a_k = \frac{2}{T} \langle \rho, \varphi_{2km} \rangle, \quad b_k = \frac{2}{T} \langle \rho, \varphi_{2km-1} \rangle \quad k \geq 1$$

les estimateurs linéaires et non biaisés de a_k et b_k sont respectivement :

$$\hat{a}_k = \frac{2}{T} \int_0^T \cos \frac{2\pi mkt}{T} dx_t \quad \text{et} \quad \hat{b}_k = \frac{2}{T} \int_0^T \sin \frac{2\pi mkt}{T} dx_t$$

En particulier, l'estimateur linéaire et non biaisé de variance minimum de $\rho(t_0)$ est

$$\begin{aligned} \hat{\rho}(t_0) = \frac{1}{\pi m} \sum_{k=1}^{+\infty} \frac{1}{k} & \left[\sin \frac{2\pi mkt_0}{T} \int_0^T \cos \frac{2\pi mku}{T} dx_u \right. \\ & \left. + \left(1 - \cos \frac{2\pi mkt_0}{T} \right) \int_0^T \sin \frac{2\pi mku}{T} dx_u \right] \end{aligned}$$

B - Problème de filtrage

Envisageons le problème suivant :

Soit y une variable aléatoire sur $(\Omega, \mathcal{F})$ telle que la variance de y et sa covariance avec chaque x_t ($t \in T$) sous la loi P_ρ , soient indépendantes de ρ (la covariance de y avec n'importe quel élément de Z est alors aussi indépendante de ρ). Nous supposerons que cette variance et ces covariances sont données ainsi que la fonction $\rho \rightarrow E_\rho(y)$ sur M .

Cherchons à l'aide d'un estimateur linéaire à partir de $(x_t, t \in T)$ à approcher "au mieux" y .

Soient

Ψ la fonction de $\Phi(R)$ telle que $\Psi(t) = \text{cov}(x_t, y)$

$$y_0 = I^{-1}(\Psi) \quad , \quad y_1 = y - y_0$$

et écrivons

$$\begin{aligned} y_0 &= y'_0 + y''_0 & y'_0 &\in V \quad , \quad y''_0 \perp V \\ \text{et pour } z \in Z, \quad z &= z' + z'' & z' &\in V \quad , \quad z'' \perp V \end{aligned}$$

On a pour $z \in Z$:

$$E_\rho(|y - z|^2) = \text{var}(y - z) + |E_\rho(y) - E_\rho(z)|^2$$

en tenant compte que, sous la loi P_ρ , $\rho \in M$, la covariance de y_1 avec n'importe quel élément de Z est nulle, on a :

$$\text{var}(y - z) = \text{var}(y_0 - z) + \text{var}(y_1).$$

D'où

$$E_{\rho} (|y - z|^2) = \text{var}(y_1) + \text{var}(y_0' - z') + \text{var}(y_0'' - z'') \\ + |E_{\rho}(y) - E_{\rho}(z')|^2$$

Et l'on peut énoncer le résultat suivant :

Théorème II.4 : "Si z est un estimateur linéaire de y , n'appartenant

" pas au sous-espace de Z , $y_0'' + V$, alors z_1 donné par

$$z_1 = y_0'' + z'$$

" est uniformément meilleur que z au sens des moindres carrés,

" c'est à dire :

$$E_{\rho} (|y - z_1|^2) < E_{\rho} (|y - z|^2) \quad \text{pour tout } \rho \in M$$

Examinons en particulier le cas des estimateurs non biaisés de y (z est un estimateur non biaisé de y si

$$E_{\rho}(y) = E_{\rho}(z) \text{ pour tout } \rho \in M).$$

Théorème II.5 : "Une condition nécessaire et suffisante pour

" qu'il existe des estimateurs linéaires non biaisés de y est

" que $h(\rho) = E_{\rho}(y)$ soit la restriction sur M d'une forme

" linéaire et continue sur V' .

" Alors, pour tout estimateur linéaire et non biaisé z de y ,

$$E_{\rho} (|y - z|^2) \text{ est indépendant de } \rho$$

" et est minimum si et seulement si

$$z = y_0'' + \hat{h}$$

" où $\hat{h}$ est l'estimateur linéaire et non biaisé de variance

" minimum de $h(\rho)$.

Ceci se déduit immédiatement de ce qui précède et du paragraphe A.

III - CAS GAUSSIEN

On se place ici dans le cas où le processus $\{x_t, t \in T\}$ a une loi temporelle de Laplace-Gauss pour chaque $P_\rho, \rho \in M$.

A - Résultats préliminaires

R. P. 1 : "Toute famille d'éléments de Z a une loi temporelle de Laplace-Gauss pour chaque $P_\rho, \rho \in M$.

Ceci résulte de la stabilité de la loi de Laplace-Gauss pour une transformation linéaire et une limite en moyenne quadratique.

R. P. 2 : "Toute famille d'éléments de l'orthogonal $V^\perp$ de V dans Z a une loi temporelle de Laplace-Gauss indépendante de ρ .

Si $x \in V^\perp$ $E_\rho(x) = (x, v_\rho) = 0$ pour $\rho \in M$.

Par ailleurs, $V^\perp \subset Z$ implique $\text{var } x$ et $\text{cov}(x, y)$ indépendantes de ρ pour x et $y \in V^\perp$ (cf. I.5').

R. P. 3 : "Si (x_i) et (y_k) sont deux familles finies de V et $V^\perp$ respectivement, alors (x_i) et (y_k) sont indépendantes pour chaque $P_\rho, \rho \in M$.

En effet, $x_i \in V$ et $y_k \in V^\perp$ implique
 $\text{cov}(x_i, y_k) = (x_i, y_k) = 0$

Nous nous proposons d'établir que la famille $(v_\rho)_{\rho \in M}$ est une statistique exhaustive pour $(P_\rho)_{\rho \in M}$.

Comme nous ne supposons connue que la loi temporelle de $\{x_t, t \in T\}$, nous poserons la définition suivante :

Définition III. 1 : "Un système de variables aléatoires $(y, y', \dots)$ définies sur $(\Omega, \mathcal{F}, P_\rho)$ est une statistique exhaustive de $\{x_t, t \in T\}$ pour $(P_\rho)_{\rho \in M}$ si :

- " - $\mathcal{A}$ (resp. $\mathcal{F}_1$) désignant la sous σ -algèbre de $\mathcal{F}$ engendrée
- " par $(y, y', \dots)$ (resp. $\{x_t, t \in T\}$).
- " - $P_\rho^{\mathcal{A}}$ désignant la probabilité conditionnelle de $F \in \mathcal{F}_1$
- " par rapport à $\mathcal{A}$, il existe une fonction $f(\omega, F)$ définie sur
- " $\Omega \times \mathcal{F}_1$, telle que pour tout $F \in \mathcal{F}_1$, $\omega \rightarrow f(\omega, F)$
- " appartienne à la classe d'équivalence de $P_\rho^{\mathcal{A}}(F)$ pour tout
- " ρ de M .

Définition III. 2 : "On dira qu'une famille d'éléments de Z est

- " une statistique exhaustive pour $(P_\rho)_{\rho \in M}$ si tout système
- " de représentants des éléments de cette famille. possède cette
- " propriété.

Théorème III. 1 : "La famille $(v_\rho)_{\rho \in M}$ constitue une statistique

" exhaustive de $\{x_t, t \in T\}$ pour $(P_\rho)_{\rho \in M}$.

Considérons un système $(u_\rho)_{\rho \in M}$ de représentants de $(v_\rho)_{\rho \in M}$.

Lemme III. 1 : "Il existe une famille de variables aléatoires

- " $\{q_t, t \in T\}$ telle que :
- " 1°) Pour tout $t \in T$, q_t représente un élément de $V^\perp$
- " 2°) La σ -algèbre G engendrée par $(u_\rho)_{\rho \in M}$ et $(q_t)_{t \in T}$
- " contient $\mathcal{F}_1$.

Projetons la classe $\tilde{x}_t$ de x_t sur V

$$\tilde{x}_t = v_t + \tilde{q}_t \quad v_t \in V, \quad \tilde{q}_t \in V^\perp \quad (\text{III.1})$$

Choisissons un représentant u_t de v_t qui soit mesurable par rapport à $\mathcal{A}$ σ -algèbre engendrée par $(u_\rho)_{\rho \in M}$ et définissons q_t par

$$q_t = x_t - u_t \quad (\text{III.2})$$

q_t représente $\tilde{q}_t$ et $x_t = u_t + q_t$ est G -mesurable donc $\mathcal{F}_1 \subset G$.

Conséquence : "Le système $\{u_\rho, \rho \in M; q_t, t \in T\}$ est une
" statistique exhaustive pour $(P_\rho)_{\rho \in M}$.

Désignons par

- $\mathcal{B}_M, \mathcal{B}_T, \mathcal{B}_M \times \mathcal{B}_T$ les σ -algèbres des boréliens de R^M, R^T et $R^M \times R^T$ respectivement.

- Y la transformation mesurable de $(\Omega, \mathcal{G})$ dans $(R^M \times R^T, \mathcal{B}_M \times \mathcal{B}_T)$ définie par :

$$Y(\omega) = (u_\rho(\omega), \rho \in M; q_t(\omega), t \in T) \quad (\text{III.3})$$

- μ_ρ la probabilité définie sur $(R^M \times R^T, \mathcal{B}_M \times \mathcal{B}_T)$
par $\mu_\rho = P_\rho Y^{-1}$.

- ν_ρ la probabilité définie sur $(R^M, \mathcal{B}_M)$ par
 $\nu_\rho(B_M) = P_\rho[Y^{-1}(B_M \times R^T)] = \mu_\rho(B_M \times R^T)$ pour $B_M \in \mathcal{B}_M$.

- λ la probabilité définie sur $(R^T, \mathcal{B}_T)$ par
 $\lambda(B_T) = P_\rho[Y^{-1}(R^M \times B_T)] = \mu_\rho(R^M \times B_T)$ pour $B_T \in \mathcal{B}_T$.

D'après R. P. 2, λ est indépendante de ρ .

Lemme III. 2 : "L'espace probabilisé $(R^M \times R^T, \mathcal{B}_M \times \mathcal{B}_T, \mu_\rho)$ est
" le produit des espace probabilisés $(R^M, \mathcal{B}_M, \nu_\rho)$ et
" $(R^T, \mathcal{B}_T, \lambda)$.

Soit d'abord $C_M \times C_T$ un rectangle mesurable dont les
"côtés" C_M et C_T appartiennent respectivement aux algèbres U_M et
 U_T des cylindres à base finie de R^M et R^T . Désignons par B_p
(resp. B_k) le borélien de R^p (resp. R^k) base de C_M (resp. C_T).

$$\mu_\rho(C_M \times C_T) = P_\rho\{\omega : (u_{\rho_1}(\omega) \dots u_{\rho_p}(\omega)) \in B_p \text{ et } (q_{t_1}(\omega) \dots q_{t_k}(\omega)) \in B_k\}.$$

D'après R. P. 3, les u_{ρ_i} et les q_{t_j} sont indépendantes,
donc :

$$\begin{aligned} \mu_\rho(C_M \times C_T) &= P_\rho \{ \omega : (u_{\rho_1}(\omega) \dots u_{\rho_p}(\omega)) \in B_p \} \\ &\times P_\rho \{ \omega : (q_{t_1}(\omega) \dots q_{t_k}(\omega)) \in B_k \} = \nu_\rho(C_M) \times \lambda(C_T) \end{aligned}$$

Pour C_M fixé, $\mu_\rho(C_M \times B_T)$ et $\nu_\rho(C_M) \cdot \lambda(B_T)$ sont deux mesures sur $(R^T, \mathcal{B}_T)$ coïncidant sur l'algèbre U_T . Elles coïncident donc sur $\sigma(U_T) = \mathcal{B}_T$.

$$\mu_\rho(C_M \times B_T) = \nu_\rho(C_M) \cdot \lambda(B_T) \text{ pour } C_M \in U_M \text{ et } B_T \in \mathcal{B}_T.$$

Alors pour B_T fixé, $\mu_\rho(B_M \times B_T)$ et $\nu_\rho(B_M) \cdot \lambda(B_T)$ sont deux mesures coïncidant sur U_M , donc sur $\sigma(U_M) = \mathcal{B}_M$.

Ainsi pour tout rectangle mesurable $B_M \times B_T$ de $R^M \times R^T$, on a : $\mu_\rho(B_M \times B_T) = \nu_\rho(B_M) \lambda(B_T)$ (III.4).

Soient :

- $A \in \mathcal{A}$ et $B_M \in \mathcal{B}_M$ tels que $A = Y^{-1}(B_M \times R^T)$
- $F \in \mathcal{F}_1$ et $E \in \mathcal{B}_M \times \mathcal{B}_T$ tels que $F = Y^{-1}(E)$

$$\begin{aligned} P_\rho(A \cap F) &= \mu_\rho[(B_M \times R^T) \cap E] \\ &= \int_{R^M \times R^T} \chi_{B_M \times R^T}(u, q) \chi_E(u, q) d\mu_\rho(u, q) \end{aligned}$$

Appliquons le théorème de Fubini et remarquons que

$$\begin{aligned} \chi_{B_M \times R^T}(u, q) &= \chi_{B_M}(u) \\ P_\rho(A \cap F) &= \int_{R^M} \chi_{B_M}(u) \left[\int_{R^T} \chi_E(u, q) \lambda(dq) \right] \nu_\rho(du) \\ &= \int_{B_M} \left[\int_{R^T} \chi_E(u, q) \lambda(dq) \right] \nu_\rho(du) \quad (\text{III.5}) \end{aligned}$$

et $g_E(u) = \int_{R^T} \chi_E(u, q) \lambda(dq)$ est $\mathcal{B}_M$ -mesurable. On remarque qu'en outre elle ne dépend pas de ρ .

Considérons la transformation mesurable S de $(\Omega, \mathcal{A})$ dans $(R^M, \mathcal{B}_M)$ telle que $S(\omega) = (u_\rho(\omega))$ $\rho \in M$.

$$\text{On a } S^{-1}(B_M) = A \text{ et } \nu_\rho = P_\rho S^{-1} \quad (\text{III.6})$$

Effectuons dans (III.5) le changement de variable indiqué dans (III.6) :

$$P_\rho(A \cap F) = \int_A g_E(S(\omega)) P_\rho(d\omega) \text{ pour } A \in \mathcal{A} \quad (\text{III.7})$$

$f(\omega, F) = g_E(S(\omega))$ est une fonction $\mathcal{A}$ -mesurable de ω et elle ne dépend pas de ρ . D'après (III.7) c'est une version de $P_\rho^{\mathcal{A}}(F)$.

On peut préciser en outre que pour tout ω , $f(\omega, F)$ est une vraie probabilité sur $\mathcal{F}_1$.

Remarque : Si l'on introduit sur $(\Omega, \mathcal{F}_1)$ la mesure P_0 telle que $\{x_t, t \in T\}$ soit gaussien centré avec $E_0(x_t \bar{x}_s) = R(t, s)$ on a pour $t \in T$: $\int x_t \nu_\rho dP_0 = \int x_t dP_\rho$.

En désignant par P'_ρ la restriction de P_ρ à $\mathcal{F}_1$ et en appliquant les résultats contenus dans [5], on a

$$\log \frac{P'_\rho(d\omega)}{P_0(d\omega)} = \nu_\rho(\omega) - \frac{1}{2} \langle \rho, \rho \rangle$$

$$\text{donc } \frac{P'_\rho(d\omega)}{P_0(d\omega)} = \exp\left(-\frac{1}{2} \langle \rho, \rho \rangle\right) \exp(\nu_\rho(\omega))$$

Or la fonction $\exp(-\frac{1}{2} \langle \rho, \rho \rangle) \exp(\nu_\rho(\omega))$ est mesurable par rapport à la σ -algèbre engendrée par $(\nu_\rho)_{\rho \in M}$.

En appliquant alors les résultats de HALMOS et SAVAGE :

An. Math. Stat. Vol. 20 (1949) p. 225 à 241, on obtient immédiatement que $(\nu_\rho)_{\rho \in M}$ est une statistique exhaustive pour $(P_\rho)_{\rho \in M}$. Toutefois la démonstration précédente permet d'affirmer en outre que la probabilité conditionnelle par rapport aux $(\nu_\rho)_{\rho \in M}$ est régulière.

BIBLIOGRAPHIE

- [1] J. HAJEK : On linear estimation theory for an infinite number of observations. Theory of probability and its applications. Vol. VI, n° 2, 1961, p 166 à 177.
- [2] J. HAJEK : Linear statistical problems in stochastic processes. Czech. Mat. Journal, T. 12, n° 3, 1962, p 404 à 444.
- [3] H.B. MANN : A theory of estimation for the fundamental random process and the Ornstein Uhlenbeck process. Sankhya, Vol. 13, 1954, p 325 à 350.
- [4] J. NEVEU : Bases mathématiques du calcul des probabilités. Masson 1964.
- [5] Yu ROZANOV : On the density of one Gaussian mesure with respect to another. Theory of probability and its applications. 7, 1962, p 82 à 87.
- [6] J. FELDMANN : Equivalence and perpendicularity of Gaussian process. Pacific J. Math., 8, 1958, p 699 à 709.