STATISTIQUE. — *Éprouver une rupture de stationnarité dans une série chronologique*, Note (*) de **Dominique Picard,** présentée par Robert Fortet.

On présente ici un test permettant d'éprouver la stationnarité d'un processus gaussien. Ce test, construit par analogie avec une généralisation multi-indice du test de Kolmogorov-Smirnov est particulièrement adapté à déceler des ruptures apparaissant sur le spectre d'une série chronologique et il possède des propriétés de robustesse évidentes.

Nous donnons principalement un théorème d'Invariance sur la fonction de répartition spectrale empirique qui permet de donner des évaluations asymptotiques de son niveau et de sa puissance dans les cas classiques de ruptures.

STATISTICS. — A Test for Detecting a Change-Point in Time Series.

Our purpose is to present a test in order to determine whether a Gaussian Process is stationary or not. This test is constructed by analogy with a multi-indiced generalization of the Kolmogorov-Smirnov test and is particularly suitable for detecting a change point in a time series occurring on the spectrum. An Invariance theorem on empirical spectral cumulative distribution function is given in order to estimate the level of the test and its powerfunction for classical cases of change point models.

dex.

I. Introduction. Notations. — Un problème fréquemment posé dans la pratique (détection de pannes, lecture d'images...) consiste au vu d'une série chronologique de détecter la présence éventuelle d'une rupture apparaissant non pas dans la moyenne mais sur le spectre de la série. La moyenne étant supposée constante, on la supposera connue quitte à traiter la série différentiée, ce qui n'affecte pas les résultats qui suivent. Comme il peut être important de déceler une rupture qui ne laisse pas constant l'ordre du modèle, nous nous placerons ici dans un cadre non paramétrique.

Or, on a vu ([2], [3]) que, pour tester l'existence d'une rupture dans une suite de n observations, réelles, indépendantes, $X_1, \ldots, X_n$, il peut être intéressant dans un cadre non paramétrique d'utiliser le test suivant qui est une généralisation en double indice du test de Kolmogorov-Smirnov :

$$(1) \qquad \sup_{x\in\mathbb{R}} \sup_{1\leqq k\leqq n-1} \frac{k}{n}\left(1-\frac{k}{n}\right)\sqrt{n}\left|\frac{1}{k}\sum_{i=1}^{k} 1_{\{X_i\leqq x\}}-\frac{1}{n-k}\sum_{i=k+1}^{n} 1_{\{X_i\leqq x\}}\right| > C,$$

pour lequel le niveau peut être approximé grâce à un théorème d'Invariance sur la fonction de répartition empirique. Nous allons ici construire un test équivalent où l'on utilisera la fonction de répartition spectrale empirique à la place de la fonction de répartition empirique.

Soit $\{X_k\}$ (k appartenant à l'ensemble $\{1, \ldots, T\}$) un processus gaussien réel centré.

Soit $\mathcal{G}$ l'ensemble des processus gaussiens réels centrés stationnaires dont la mesure spectrale est absolument continue par rapport à la mesure de Lebesgue sur $[-\pi, \pi]$ et dont la densité spectrale f vérifie les deux propriétés suivantes :

$$(1) \qquad F(\lambda)=\int_0^\lambda f(u)\,du \text{ est strictement croissante,}$$

$$(2) \qquad \text{il existe } \delta>0 \quad \text{tel que} \quad \int_0^\pi f^{2+\delta}(u)\,du<+\infty.$$

On éprouvera donc l'hypothèse $H_0 : \{\{X_k\}\in\mathcal{G}\}$, plus particulièrement contre l'hypothèse :

$$H_1 : \quad \{\exists k_0,\ 1<k_0<T,\ X_k=Y_k\,1_{k<k_0}+Y'_k\,1_{k\geqq k_0},$$
$$\forall k\in\{1, \ldots, T\},\ \{Y_k\}\in\mathcal{G},\ \{Y'_k\}\in\mathcal{G}, F_y\neq F_{y'}\},$$

où F_y (resp. $F_{y'}$) désigne ici la répartition spectrale du processus $\{Y_k\}$ (resp. $\{Y'_k\}$).

Définissons pour $1 \leqq k < T$:

$$I_k(\lambda) = \frac{1}{2\pi k} \left| \sum_{j=1}^{k} X_j e^{-i\lambda j} \right|^2,$$

$$I^*_{T-k}(\lambda) = \frac{1}{2\pi(T-k)} \left| \sum_{j=k+1}^{T} X_j e^{-i\lambda j} \right|^2,$$

$$\hat{F}_k(\lambda) = \int_0^\lambda I_k(u)\,du, \qquad \hat{F}^*_{T-k}(\lambda) = \int_0^\lambda I^*_{T-k}(u)\,du,$$

$$Z_T\left(\lambda, \frac{k}{T}\right) = \frac{k}{\sqrt{T}}(\hat{F}_k(\lambda) - F(\lambda)),$$

$$Z_T = \operatorname*{Sup}_{\lambda \in [0,\pi]} \operatorname*{Sup}_{k \in \{1, \ldots, T-1\}} \sqrt{T}\frac{k}{T}\left(1 - \frac{k}{T}\right) |\hat{F}_k(\lambda) - \hat{F}^*_{T-k}(\lambda)|.$$

II.4. Résultats principaux

Théorème 1. — *Le test* φ_T *de région de rejet* $\{ Z_T > C\sqrt{2\pi G_2} \}$ *a un niveau approximable pour* T *grand par la quantité :*

$$P\{ \operatorname*{Sup}_{t \in [0,1]} \operatorname*{Sup}_{\lambda \in [0,1]} |B(\lambda, t)| > C \},$$

où :

$$G_2 = \int_0^\pi f^2(l)\,dl$$

et $B(\lambda, t)$ *est le processus gaussien centré de* $\mathscr{C}([0, 1]^2)$ *de covariance :*

$$EB(\lambda, t)B(\mu, s) = [\inf(\lambda, \mu) - \lambda\mu] \times [\inf(t, s) - st].$$

Le test φ_T *a donc un niveau facilement approximable : en effet, le processus* B *est classique : c'est le même processus que l'on obtient pour* [1]. *Nous l'avons tabulé à l'aide de méthodes de Monte-Carlo. En particulier, aux valeurs de* $\alpha = 0{,}05$, $0{,}025$, $0{,}10$ *correspondent avec une bonne approximation les constantes* $C_\alpha = 0{,}775\,0$, $0{,}814\,6$, $0{,}712\,7$.

Ce théorème est en fait un corollaire du principe d'Invariance suivant qui est également l'outil essentiel pour les calculs de puissance, du moins pour ce qui concerne les points de la contre hypothèse appartenant à H_1. (Il permet en particulier de montrer que ce test est consistant pour toute contre-hypothèse du type $H_1(k_0 = T\,t_0)$, $t_0 \in]0, 1[$.)

Théorème 2. — *Sous l'hypothèse* H_0, *la suite de processus* $Z_T(.,.)$ *définie par :* $\forall \lambda \in [0, \pi]$, $Z_T(\lambda, .)$ *est la ligne polygonale joignant les points* $(0, 0)$ *et :*

$$\left(\frac{k}{T}, Z_T\left(\lambda, \frac{k}{T}\right)\right), \qquad k = 1, \ldots, T,$$

converge étroitement dans $\mathscr{C}([0, \pi] \times [0, 1])$ *vers le processus* $Z(.,.)$ *gaussien centré de covariance :*

$$EZ(\lambda, t)Z(\mu, s) = \int_0^{\inf(\lambda, \mu)} f^2(l)\,dl \times \inf(t, s).$$

Schéma de démonstration. — Elle utilise fondamentalement des évaluations et majorations de moments qui s'inspireront des résultats de [5].

1. La convergence des suites de marginales finies est classique.

2. La tension dans $\mathscr{C}([0, \pi] \times [0, 1])$ de la suite $Z_T(.,\ .)$ est une conséquence en utilisant le classique lemme de Csensov (*cf.* [1]) du résultat suivant :

Il existe $r>0$, $\gamma>0$, $\alpha>1$ tels que :

$$\forall T \in \mathbb{N}^*, \quad \forall \lambda \in [0, \pi], \quad \forall \mu < \lambda, \quad \forall t \in [0, 1], \quad \forall s < t$$
$$E(Z_T(\lambda, t) - Z_T(\mu, t) - Z_T(\lambda, s) + Z_T(\mu, s))^r \leqq \gamma [(t-s)(\lambda-\mu)]^\alpha.$$

On trouvera les détails de cette démonstration dans [8].

III. Commentaires. — Il reste, dans le cas général pour pouvoir utiliser le test φ_T à remplacer G_2 par une estimation consistante pour que le théorème I reste inchangé. Ceci peut être fait d'après [5], par exemple avec :

$$\hat{G}_2 = \frac{1}{4\pi} \sum_{j=0}^{T^{1/2}} \left(\frac{1}{T} \sum_{k=0}^{T-j} X_k X_{k+j} \right)^2.$$

Nous avons traité ici le cas où l'observation $\{X_k\}$ est en temps discret. La transposition en temps continu ne présente pas de difficulté.

Si l'intérêt principal de ce test est son caractère non paramétrique, en revanche il présente l'inconvénient de ne pas conduire à une estimation de l'instant de rupture quand il existe. L'une des raisons en est la pondération $k/T(1-(k/T))$ introduite dans la statistique de test. Il est nécessaire d'introduire une pondération : en effet : (la limite s'entend en probabilité sous H_0) :

$$\lim_{T \to \infty} \sup_{\lambda \in [0, \pi]} \sup_{k \in \{1, \ldots, T-1\}} \sqrt{T} \, |\hat{F}_k(\lambda) - \hat{F}^*_{T-k}(\lambda)| = +\infty.$$

En revanche, pour tout $\varepsilon>0$,

$$\sup_{\lambda \in [0, \pi]} \sup_{k \in \{1, \ldots, T-1\}} \sqrt{T} \left(\frac{k}{T} \left(1 - \frac{k}{T} \right) \right)^{1/2+\varepsilon} |\hat{F}_k(\lambda) - \hat{F}^*_{T-k}(\lambda)|,$$

a une limite finie en loi sous H_0, de sorte que de nombreuses pondérations sont envisageables et permettent de construire un test. Toutefois, plus ε est grand plus le test a tendance à pénaliser fortement les bords et fait perdre de la puissance sur des contre-hypothèses du type $H_1(k_0)$ avec k_0 petit ou près de T. La loi limite obtenue avec $\varepsilon = 1/2$ présente l'avantage d'être de plus homogène par rapport au couple (λ, t).

Enfin, on peut remarquer que le théorème 2 permet d'envisager d'autres procédures que celle adaptée de Kolmogorov-Smirnov, par exemple un test du type Cramer-Von-Mises fondé sur la statistique :

$$\frac{1}{\sqrt{T}} \sum_{k=2}^{T-1} \left(\frac{k}{T} \left(1 - \frac{k}{T} \right) \right)^2 \int_0^1 (\hat{F}_k(\lambda) - \hat{F}^*_{T-k}(\lambda))^2 \, d\lambda,$$

présente l'avantage d'avoir une loi limite proportionnelle à $\int_0^1 \int_0^1 B^2(\lambda, t) \, dt \, d\lambda$ qui est connue et tabulée (*cf.* [7]).

(*) Reçue le 15 juillet 1982, acceptée le 11 octobre 1982.

[1] N. N. CSENSOV, *Proceedings all Union Conf. Theory Prob. and Math. Statistics*, Erevan, 1955, Izda. Akad. Nank., Aronjan, S.S.R., Erevan, 1960, p. 280-285.

[2] J. DESHAYES et D. PICARD, *Comptes rendus*, 292, série I, 1981, p. 449.

[3] J. DESHAYES et D. PICARD, *Testing for a Change-Point in Statistical Models*, Rapport technique, 1981, Université Paris-Sud, Orsay.

[4] J. DESHAYES et D. PICARD, *Principe d'Invariance sur le Processus de Vraisemblance*, 1982, Rapport Technique, Université Paris-Sud, Orsay.

[5] I. A. IBRAGIMOV, *Theory Prob. and Appl.*, 1963, p. 366-400.

[6] KOMLOS MAJOR TUSNADY, *An Approximation of Partial Sums of Independant rv's and the sample DF:* IZ wahrscheinlichkeitstheorie-Verw: I, Gebiete 32, 1975, p. 111-131; II, Gebiete 34, 1976 p. 33-58.

[7] BLUM-KIEFER-ROSENBLATT, *Ann. Math. Stat.*, 1961, p. 485-497.

[8] D. PICARD, *Rupture dans une série chronologique. Test et Estimation*, 1982, Rapport technique, Université Paris-Sud Orsay.

Université Paris-Sud, Département de Mathématiques n° 425,
E.R.A. n° 532, *Statistique appliquée*, 91405 *Orsay*.