ANNALES SCIENTIFIQUES
DE L'UNIVERSITÉ DE CLERMONT-FERRAND 2
Série Mathématiques

D. GUEGAN

Comparaison de tests d'hypothèses pour des processus gaussiens

Annales scientifiques de l'Université de Clermont-Ferrand 2, tome 69, série *Mathématiques*, n° 19 (1981), p. 57-65

<http://www.numdam.org/item?id=ASCFM_1981__69_19_57_0>

COMPARAISON DE TESTS D'HYPOTHESES POUR DES PROCESSUS GAUSSIENS

D. GUEGAN

Université de Paris-Nord

On se propose d'étudier des tests du maximum de vraisemblance qui sont une généralisation des tests de Neymann-Pearson. Ce sont des tests sup/sup. On se place d'un point de vue asymptotique et non local.

On va tout d'abord présenter un test étudié dans le cas des hypothèses séparées asymptotiquement, test que l'on comparera ensuite à des tests étudiés dans le cas d'hypothèses emboitées ou contigües.

Test d'hypothèses séparées - Présentation du problème

Soit un processus stationnaire du second ordre $(X_t)_{t \in \mathbb{R}}$, supposons observés $X_1, \ldots, X_N$, et soit à tester Θ_0 contre Θ_1, hypothèses composées. En un point $\alpha \in \Theta_0$, la densité spectrale du processus est notée $f(\lambda,\alpha)$, en un point $\beta \in \Theta_1$, $g(\lambda,\beta)$. Il leur correspond une famille de distributions notée P_o sous $Ⓗ_o$ et P_1 sous $Ⓗ_1$, définies sur un espace approprié d'échantillons satisfaisant les hypothèses précisées ci-dessous :

L'espace des paramètres est supposé compact.

On note f', g', ... les dérivées des fonctions prises par rapport à λ.

(1) $f(\lambda,\alpha)$, $f'(\lambda,\alpha)$, $f''(\lambda,\alpha)$ appartiennent à $L^2(d\lambda)$.

(2) $f(.,\alpha)$ est continûment dérivable jusqu'à l'ordre 2.

(3) $\dfrac{g(\lambda,\beta)}{f^2(\lambda,\alpha)}$ est intégrable.

Soit $b(\lambda,\alpha) = f^{-1}(\lambda,\alpha)$, on suppose que ses dérivées première, seconde et troisième par rapport à λ existent :

(4) $b'(\lambda,\alpha)$ est absolument continue en (λ,α).

(5) $b''(\lambda,\alpha)$ et $b'''(\lambda,\alpha)$ sont continues en (λ,α).

(6) m_1 et m_2 sont des constantes telles que $\forall\alpha$, $\forall\lambda$

$0 \leqslant m_1 \leqslant f(\lambda,\alpha)$ et $|f''(\lambda,\alpha)| \leqslant m_2$

On note $L_N(X,\alpha)$ le log vraisemblance du processus sous l'hypothèses $Ⓗ_o$ et $L_N^*(X,\beta)$ sa log vraisemblance sous $Ⓗ_1$, alors sous les hypothèses (1) - (6), si $I_N(\lambda)$ représente le périodogramme du processus, on a l'approximation suivante du processus dû à Whittle cf. [3] :

Sous $Ⓗ_o$ $L_N(X,\alpha) = \frac{N}{2\pi} \left[\int \log f(\lambda,\alpha)d\lambda + \dfrac{I_N(\lambda)}{f(\lambda,\alpha)} d\lambda \right]$.

Sous $Ⓗ_1$ $L_N^*(X,\beta) = \frac{N}{2\pi} \left[\int \log g(\lambda,\beta)d\lambda + \dfrac{I_N(\lambda)}{g(\lambda,\beta)} d\lambda \right]$.

On note $\hat{\alpha}_N$ et $\hat{\beta}_N$, les estimateurs du maximum de vraisemblance supposés existés et uniques, et soit β_α le point de $Ⓗ_1$ le plus proche de $Ⓗ_o$, il est défini par la relation :

$$\int \frac{f(\lambda,\alpha)}{g(\lambda,\beta_\alpha)} g'_\beta(\lambda,\beta_\alpha)\, d\lambda = 0.$$

On a alors le théorème suivant dont on trouvera une démonstration dans [6].

<u>Théorème 1</u>

Sous les hypothèses (1) - (2) $\hat{\alpha}_N \to \alpha$ P_o p.s.

Sous les hypothèses (1) - (4) $\hat{\beta}_N \to \beta_\alpha$ P_o p.s.

Sous les hypothèses (1) - (5) $\hat{\alpha}_N$ et $\hat{\beta}_N$ sont des estimateurs asymptotiquement normaux sous Ⓗ$_o$ et Ⓗ$_1$.

On fait l'hypothèse supplémentaire suivante :

(7) On se fixe deux voisinages de α et de β_α tels que :

$P_o(\hat{\alpha}_N \notin V_\alpha) = 0(1)$, $P_1(\hat{\beta}_\alpha \notin V_s) = 0(1)$, alors les fonctions

$$\frac{f'(\lambda,\alpha')}{f(\lambda,\alpha')}, \quad \frac{f'(\lambda,\alpha')g'(\lambda,\beta')}{g'(\lambda,\beta')} \quad f''(\lambda,\alpha')h(\lambda) \text{ où } h(\lambda) = \frac{1}{f(\lambda,\alpha')} - \frac{1}{g(\lambda,\beta')}$$

sont majorées par des fonctions de $L^2(d\lambda)$ pour tout $\alpha' \in V_\alpha$ et $\beta' \in V_{\beta_\alpha}$.

Si on note $f^N(X,\alpha)$ et $g^N(X,\beta)$ le vraisemblance du processus sous Ⓗ$_o$ et Ⓗ$_1$, on définit l'information de Kullback $K(\alpha,\beta)$de la manière suivante : étant donné la suite

$K_N(\alpha,\beta) = E_\alpha \log \dfrac{f^N(X,\alpha)}{g^N(X,\beta)}$, soit $K(\alpha,\beta) = \lim\limits_N \frac{1}{N} K_N(\alpha,\beta)$, et on montre

(cf. [6]) que sous les hypothèses (1) - (6), pour les processus, elle vaut :

$$K(\alpha,\beta) = \frac{1}{2\pi}\left[\int \log \frac{f(\lambda,\alpha)}{g(\lambda,\beta)} d\lambda + 1 - \int \frac{f(\lambda,\alpha)}{g(\lambda,\beta)} d\lambda\right].$$

Alors on dira que les hypothèses Ⓗ$_o$ et Ⓗ$_1$ sont asymptotiquement séparées si

$$J(Ⓗ_o, Ⓗ_1) = \inf_{\alpha \in Ⓗ_o,\ \beta \in Ⓗ_1} K(\alpha,\beta) > 0.$$

Nous considérons des tests basés sur la statistique

$Z_N = \dfrac{\sup_\alpha L_N(X,\alpha)}{\sup_\beta L_N(X,\beta)} = \dfrac{L_N(X,\hat{\alpha}_N)}{L_N(X,\hat{\beta}_N)}$, et pour un niveau a donné, on définit une région de rejet $\{Z_N > C_N\}$.

On note $\hat{f}$, $\hat{g}$, $\hat{h}$, $\hat{D} = \int \log \dfrac{\hat{f}(\lambda,\alpha)}{\hat{g}(\lambda,\beta)}\, d\lambda$ les fonctions prises aux points $\hat{\alpha}_N$ et $\hat{\beta}_N$.

Sous $(H)_o$, Z_N tend p.s. vers une constante du type $\frac{A}{B}$ avec $A - B = K(\alpha,\beta_\alpha)$. La distribution de Z_N est symétrique sous $(H)_o$ et $(H)_1$, donc sous $(H)_1$, Z_N tend vers la même constante positive, et les hypothèses ne se séparent pas à distance finie.

On considère alors la statistique $\hat{L} = L_N(X,\hat{\alpha}_N) - L_N(X,\hat{\beta}_N)$. Sous $(H)_o$, $\hat{L}$ se comporte comme $NK(\alpha,\beta_\alpha)$ et tend donc p.s. vers $+\infty$ et par symétrie sous $(H)_1$, $\hat{L}$ tend p.s. vers $-\infty$. Les hypothèses se séparent bien et le test basé sur cette statistique est consistant, malheureusement la statistique $\hat{L}$ n'est pas libre sous $(H)_o$ et ne permet pas de construire une région de rejet indépendante de α. On peut alors espérer avoir un comportement en loi qui soit libre en modifiant cette statistique, d'où l'idée de la centrer et de la normer. Nous allons donc étudier la distribution de la statistique suivante : $T_N = \dfrac{\hat{L} - E_{\hat{\alpha}}\hat{L}}{(\mathrm{var}_{\hat{\alpha}}\hat{L})^{1/2}}$, où $E_{\hat{\alpha}}$ et $\mathrm{var}_{\hat{\alpha}}$ représentent l'espérance et la variance prises pour $P_{\hat{\alpha}_N}$.

Etude de la distribution de T_N

Cette étude peut être trouvée en détail dans [6] et [7] . Nous en donnons les grandes lignes avant de passer aux résultats.

L'Etude repose

- sur la non corrélation asymptotique des statistiques de tests et d'estimation (cf. [6])
- le développement de Taylor à l'ordre deux des fonctions $f(\lambda,\alpha)$ et $h(\lambda)$ aux points $\hat{\alpha}_N$ et $\hat{\beta}_N$
- le théorème de Grenander et Rosenblatt (cf. [5] et [6]).

Alors :

<u>Théorème 2</u> : Sous les hypothèses (1) - (7), la distribution asymptotique de la statistique T_N sous $Ⓗ_o$ est une loi normale centrée $N(0,\Gamma(\alpha))$ avec

$$\Gamma(\alpha) = \frac{1}{4\pi^2}\,\frac{K'^2_\alpha(\alpha,\beta_\alpha)}{I(\alpha)V(\alpha)} \quad \text{où} \quad K'_\alpha(\alpha,\beta) = \int \frac{f'(\lambda,\alpha)}{g(\lambda,\beta)}\,d\lambda$$

$$I(\alpha) = \int \frac{f'^2(\lambda,\alpha)}{f^2(\lambda,\alpha)}\,d\lambda \text{ et } V(\alpha) = 2\pi\left[1 - 2\int \frac{f(\lambda,\alpha)}{g(\lambda,\beta_\alpha)}\,d\lambda + \int \frac{f^2(\lambda,\alpha)}{g^2(\lambda,\beta_\alpha)}\,d\lambda\right]$$

On remarque que la statistique obtenue pour T_N n'est pas libre, elle dépend de α. On va donc estimer $\Gamma(\alpha)$ par $\Gamma(\hat{\alpha}_N)$. $\Gamma(\alpha)$ ne fait intervenir que des contantes dépendant de l'information, obtenues par simples passages à la limite sous le signe somme.

Par exemple :

$$I(\hat{\alpha}_N) = \int \frac{f'^2(\lambda,\hat{\alpha}_N)}{f^2(\lambda,\hat{\alpha}_N)}\,d\lambda = \int \frac{f'^2(\lambda,\alpha)}{f^2(\lambda,\alpha)}\,d\lambda + (\hat{\alpha}_N - \alpha)\,l_N(\alpha) \quad \text{où}$$

$l_N(\alpha)$ est une fonction bornée et $(\hat{\alpha}_N - \alpha)$ est un $\sigma(\frac{1}{\sqrt{N}})$. Donc l'erreur commise par ce genre d'approximation est de l'ordre de $\frac{1}{\sqrt{N}}$; on a donc

$\mathrm{Var}\, T_N = \Gamma(\alpha) + 0(\frac{1}{N})$ et T_N tend vers $N(0,\Lambda_N)$ où $\Lambda_N = \Gamma(\alpha) + 0(\frac{1}{N})$.

<u>Théorème 3</u> : Sous les hypothèses (1) - (7), $T_N \times \Gamma(\hat{\alpha}_N)^{-1}$ tend vers une loi normale centrée réduite.

Et on montre sous $Ⓗ_1$ (cf. [7]), en faisant jouer un rôle symètrique aux hypothèses $Ⓗ_o$ et $Ⓗ_1$, en particulier en remarquant que

$\hat{\alpha}_N \to \alpha_\beta \; P_1$, p.s. et $\hat{\beta}_N \to \beta \; P_1$.p.s. que :

<u>Théorème 4</u> : Sous les hypothèses (1) - (7), la distribution de la statistique T_N sous l'hypothèse Ⓗ$_1$ tend vers $+\infty$.

Les théorèmes 3 et 4 permettent de construire un test consistant basé sur la statistique T_N renommée convenablement dont la distribution en loi est libre, avec une région de rejet indépendante de α, de la forme $\{T_N > C_N\}$ correspondant au niveau a défini par $P_o(\{T_N > C_N\}) = a$.

Alors la puissance de ce test est voisine de $\left(\frac{C_N - N\Gamma(\hat{\alpha}_N)}{S(\beta)}\right)$ où ϕ représente la fonction de distribution de la loi $N(0,1)$ et $S(\beta) = \frac{1}{2\pi} \frac{R(\beta)}{V(\beta)^{1/2}}$

avec $R(\beta) = \int \frac{f(\lambda,\alpha_\beta)}{g(\lambda,\beta)} d\lambda + \int \frac{g(\lambda,\beta)}{f(\lambda,\alpha_\beta)} d\lambda - 2$ et

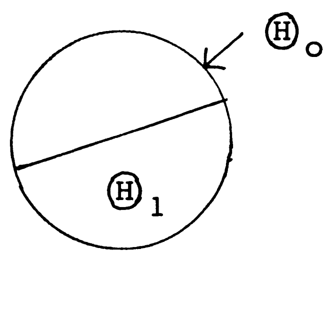
$$V(\beta) = 2\pi \left[\int \frac{g^2(\lambda,\beta)}{f^2(\lambda,\alpha_\beta)} d\lambda + 1 - \int \frac{g(\lambda,\beta)}{f(\lambda,\alpha)} d\lambda \right] .$$

<u>Autres Points de Vue</u>

Le test T_N est un test d'hypothèses séparées. Il permet de tester par exemple un processus autorégressif d'ordre 1, $X_t - \alpha X_{t-1} = \varepsilon_t$ avec $\alpha > \alpha_o \neq 0$ contre une moyenne mobile d'ordre 1, $X_t = \varepsilon_t + \beta\varepsilon_{t-1}$ avec $\beta > \beta_o \neq 0$.

La situation peut être schématisée de la manière suivante :

Mais si les deux familles ont un point commun, et en particulier dès que ce point commun représente la vraie valeur du paramètre, alors on est dans le cas des hypothèses continues. On est alors dans une situation du type suivant :

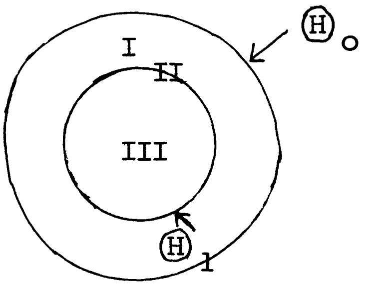


Un article qui nous semble décisif pour ce problème est celui de K.O. DZHAPARIDZE [4]. Celui-ci étudie la distribution du rapport de vraisemblance sous $\textcircled{H}_o$ et $\textcircled{H}_1$, sous les hypothèses de LE CAM [8], donc dans le cadre de la contiguïté. Il teste sous $\textcircled{H}_o$ $f(\lambda,\alpha,0)$ contre $f(\lambda,\alpha,\beta)$ sous $\textcircled{H}_1$ avec $\beta = 0(\frac{1}{\sqrt{N}})$, soit donc $\beta = 0$ contre $\beta \neq 0$. Dans ce cas DZHAPARIDZE montre alors que la distribution asymptotique de

$$-2\,\frac{L_N(\chi,\hat{\alpha},0)}{L_N(\lambda,\alpha,\frac{\hat{\beta}}{\sqrt{N}})} \quad \text{est un } \chi'^2.$$

La démonstration repose essentiellement sur les hypothèses de LE CAM, le développement de Taylor et la convergence des estimateurs à la vitesse $\frac{1}{\sqrt{N}}$. En fait en se plaçant sous l'hypothèse $\textcircled{H}_1$, il donne un résultat de puissance, et étant dans le cadre des hypothèses contigües, cela lui permet de définir une région de rejet et d'obtenir ainsi un test.

Le point de vue de DZHAPARIDZE nous semble le bon point de vue dès que l'on est amené à tester des hypothèses emboitées, ce que l'on peut schématiser par



En effet, si on doit tester un élément de la région I contre un élément de la région III on peut alors traiter ce cas par les hypothèses séparées, le seul cas posant un problème est celui de la région II, la frontière en quelque sorte et ce cas entre dans le cadre envisagé par DZHAPARIDZE.

Néanmoins on peut noter le point de vue de AKAIKE [1] repris par BOUAZIZ dans sa thèse [2] ; il étudie la distribution asymptotique du rapport de vraisemblance par des estimateurs supposés asymptotiquement normaux et convergeants vers une même valeur α. Soit

$$\begin{array}{lcl} \hat{\alpha}_N & \searrow & \\ & & \alpha \\ \hat{\beta}_N & \nearrow & \end{array}$$

Il montre alors que si p est le dimension de l'espace des paramètres Θ_α et r la dimension de l'espace des paramètres Θ_β , la loi du rapport de vraisemblance est alors sous Θ_o un $\chi^2(p-r)$.

La démonstration revient à considérer qu'il est en présence d'une différence de deux formes quadratiques qui coincident sur un ensemble commun à cause des hypothèses emboitées.

On peut donc remarquer en définitive que la situation étudiée par BOUAZIZ est envisagée par le test proposé par DZHAPARIDZE et n'apporte donc pas d'information supplémentaire, et que le test T_N se place tout à fait hors du cadre étudié par DZHAPARIDZE.

BIBLIOGRAPHIE

[1] AKAIKE, H. - "Information theory and an extension of the maximum likelihood principle"
2nd Inter. Symp. on Information Theory, Budapest Akademia Kiado, pp. 267-281, 1973.

[2] BOUAZIZ, M. - "Identification de l'ordre de dépendance dans les series temporelles"
Thèse de 3ème Cycle, Orsay, N° 2450, 1978.

[3] DACUNHA-CASTELLE, D. - "Vitesse de convergence pour certaines problèmes statistiques"
Ecole de Proba. de St. Flour, Springer, 1977.

[4] DZHAPARIDZE, K.O.- "Tests of composite hypotheses for random variables and stochastic processes"
Theory of Proba. and its Appl., Vol.22, N° 2, 1977, p. 104-120.

[5] GRENANDER, H. ROSENBLATT, M. - "Statistical spectral analysis of time series arising from stationary stochastic processes"
A.M.S., Vol. 24, p. 537-558, 1953.

[6] GUEGAN, D. - "Loi limite de la statistique de tests de vraisemblance"
Thèse de 3ème Cycle, Orsay, N° 2201, 1977.

[7] GUEGAN, D. - "Tests d'hypothèses séparés pour des processus stochastiques"
Note aux C.R. (à paraître).

[8] LE CAM, L. - "Théorie asymptotique de la décision statistique"
Montréal University Press, 1969.