

Astérisque

DIDIER DACUNHA-CASTELLE

Formule de Chernoff pour des rapports de vraisemblance

Astérisque, tome 68 (1979), p. 25-31

<http://www.numdam.org/item?id=AST_1979__68__25_0>

© Société mathématique de France, 1979, tous droits réservés.

L'accès aux archives de la collection « Astérisque » (<http://smf4.emath.fr/Publications/Asterisque/>) 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/>

FORMULE DE CHERNOFF POUR DES RAPPORTS DE VRAISEMBLANCE

D. DACUNHA-CASTELLE.

I. GÉNÉRALITÉS.

Soient P_0, P_1 deux probabilités de densité f_0 et f_1 par rapport à une mesure ν . On suppose $P_0 \neq P_1$ et P_0 équivalente à P_1 .

Soit $\Phi(t) = E_{P_0} \exp t \log \frac{f_1}{f_0}$.

Pour $t \in [0, 1]$, $\Phi(t)$ est finie, $0 \leq \Phi(t) \leq 1$ d'après l'inégalité de Holder. Si $t \geq 0$, Φ est en fait définie (et finie) sur $[0, s[$, $s = \text{ess sup} \log \frac{f_1}{f_0}$, ($s \geq 1$). En s , Φ est finie si et seulement si

$$P_0 \left(\log \frac{f_1}{f_0} = s \right) = 0.$$

On pose $h(a) = \sup_{t \geq 0} at - \log \phi(t)$

(fonction de Cramer du couple P_0, P_1), et

$$K(P_0, P_1) = E_{P_0} \log \frac{f_0}{f_1} \cong \infty$$

$$K(P_1, P_0) = - E_{P_1} \log \frac{f_0}{f_1} \cong \infty.$$

ϕ est analytique sur $]0, 1[$ et

$$\lim_{t \downarrow 0} \Phi'(t) = -K(P_0, P_1)$$

$$\lim_{t \uparrow 1} \Phi'(t) = K(P_1, P_0)$$

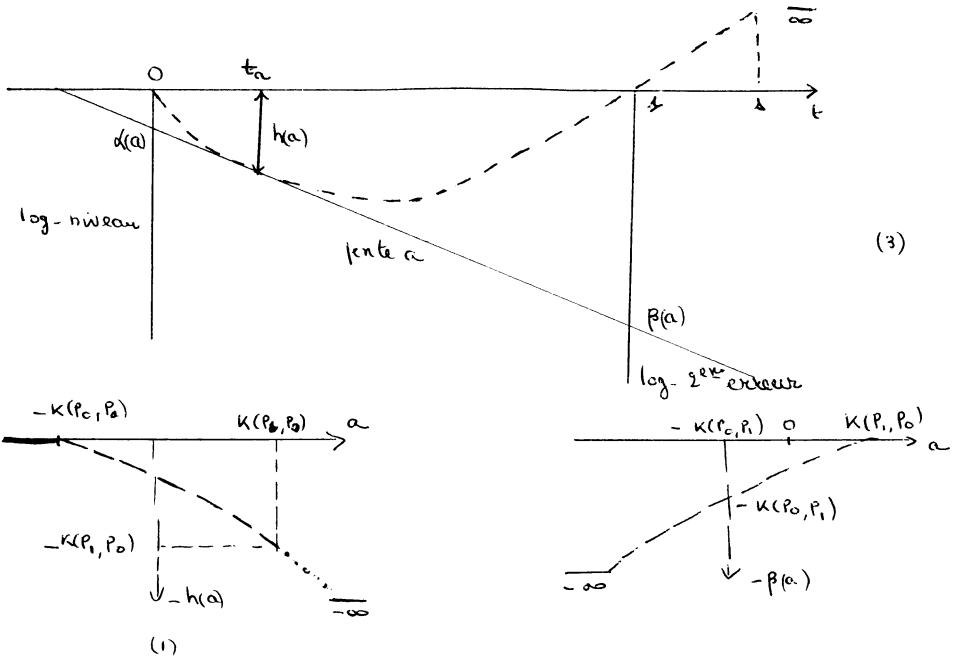
On a $h(a) \rightarrow 0$ pour $a \rightarrow -K(P_0, P_1)$, et on définit h partout dans $\bar{\mathbb{R}}$, en posant $h(a) = 0$ pour $a \leq -K(P_0, P_1)$.

h ainsi définie est convexe, strictement convexe et continue, sauf en s dans le cas où $P_0 \left(\log \frac{f_1}{f_0} = s \right) > 0$.

EXPOSÉ 2

Enfin $E_{P_1} \exp t \log \frac{f_1}{f_0} = \hat{\Phi}(1-t)$ et donc si $a \in K(P_1, P_0)$,

$$-\sup_{t > 0} (-at - \log \hat{\Phi}(1-t)) = a-h(a)$$



Soit f une fonction convexe sur \mathbb{R} à valeurs dans $\mathbb{R} \cup \{+\infty\}$.

La duale \hat{f} de f est définie par

$$\hat{f}(t) = \sup_{x \in \mathbb{R}} tx - f(x)$$

\hat{f} est évidemment convexe, à valeurs dans $\mathbb{R} \cup \{+\infty\}$. Dans le cas où f est dérivable, alors \hat{f} est dérivable, et \hat{f}' est la fonction réciproque $(f')^{-1}$ de f' .

Les résultats de convergence suivant sont faciles (pour des fonctions de variables réelles). On note $Df = \{x, f(x) < \infty\}$.

1) Soit f_n une suite de fonctions convexes à valeurs dans $\mathbb{R} \cup \{\infty\}$,

RAPPORTS DE VRAISEMBLANCE

convergeant vers f en tout point, sauf peut-être les points de $(Df)^*$, frontière de Df . Supposons $\overset{\circ}{Df} \neq \emptyset$ ($\overset{\circ}{Df}$ intérieur de Df). Alors la convergence est uniforme sur tout compact de $\overset{\circ}{Df}$ (et de $R - \overline{Df}$), (\overline{Df} fermeture de Df).

2) De plus $\hat{f}_n(x) \rightarrow \hat{f}(x)$ sauf peut-être en $x \in (Df)^*$.

3) Soit alors Z une variable aléatoire (non dégénérée). $\phi(t) = E \exp tZ$ est convexe à valeurs dans $\mathbb{R} \cup \{\infty\}$. $\log \phi(t)$ est convexe (inégalité de Hölder).

Il est immédiat de vérifier que

$$\overline{\{t, \log \phi(t) < \infty\}} = \text{support } Z$$

et de manière plus précise

si $b = \sup \{x, x \in \text{support } Z\}$, on a

$$\log \phi(b) < \infty \text{ équivaut à } P(Z = b) = 0$$

Dans la suite interviendra de manière décisive la duale de $\log \phi$.

Soit $(P_{0,n}, P_{1,n})_{n \in \mathbb{N}}$ une suite de couples de probabilités équivalentes sur \mathbb{R}^n , $n \in \mathbb{N}$. On dira que cette suite est de Chernoff si

$$\lim_{n \rightarrow \infty} \phi_n^{1/n}(t) = \exp L(t), \quad t \in [0,1]$$

où L est une fonction convexe sur $[0,1]$.

On a donc $\lim_{n \rightarrow \infty} \frac{1}{n} \log \phi_n(t) = L(t)$.

Si de plus $\frac{1}{n} \phi_n'(0)$ converge, on appellera $K(P_0, P_1)$ sa limite, qui sera définie comme l'information de Kullback de P_1 par rapport à P_0 .

Le test de vraisemblance au seuil α de P_0 contre P_1 a pour région de rejet de P_0 , $\left\{ \frac{1}{N} \log \frac{f_{1,N}}{f_{0,N}} > \alpha \right\}$ pour N fixé.

Le théorème du chapitre précédent s'écrit ici sous la forme suivante :

Théorème : Soit une suite $(P_{0,n}, P_{1,n})$ de Chernoff, supposons $K(P_0, P_1)$, $K(P_1, P_0)$ définies et $\alpha \in (-K(P_0, P_1), K(P_1, P_0))$. Alors

EXPOSÉ 2

$$\lim_{N \rightarrow \infty} \frac{1}{N} \log P_{0,N} \left(\frac{1}{N} \log \frac{f_{1,N}}{f_{0,N}} > a \right) = -h(a)$$

où h est la duale de L , $L(t) = \lim_{N \rightarrow \infty} \frac{1}{N} \log \phi_N(t)$.

Ce théorème permet de régler asymptotiquement le seuil du test à un niveau $\exp - N\alpha$. On devra avoir $h(a_N) \neq \alpha$, et l'on aura $a_N \rightarrow a$ tel que $h(a) = \alpha$.

Ce réglage étant fait, on peut calculer

$$\begin{aligned} & \lim_{N \rightarrow \infty} \frac{1}{N} \log P_{1,N} \left(\frac{1}{N} \log \frac{f_{1,N}}{f_{0,N}} \leq a \right) \\ &= \lim_{N \rightarrow \infty} \frac{1}{N} \log P_{1,N} \left(\frac{1}{N} \log \frac{f_{0,N}}{f_{1,N}} \geq a \right) \\ &= \beta \end{aligned}$$

tel que $\alpha + \beta = a$, d'après une formule vue plus haut. Le test de Neyman-Pearson au niveau $\exp - n\alpha$ est donc tel que son erreur de 2ème espèce est $\exp - \beta_n$, avec $\lim_{N \rightarrow \infty} \beta_n = a - \alpha$ avec $\alpha = h(a)$. (Voir dessin).

Il est immédiat d'appliquer ce résultat à une suite de variables indépendantes équidistribuées. Nous allons l'appliquer à un cas qui n'est pas traité complètement, dans la littérature, celui de deux processus gaussiens stationnaires.

II. CAS DE DEUX PROCESSUS GAUSSIENS (d'après [3]).

[0] se réfère au chapitre I.

Les notations sont celles de [3] qui contient une bibliographie complète. F et G sont deux mesures gaussiennes centrées stationnaires sur $\mathbb{R}^{\mathbb{Z}}$, et X^N désigne l'observation (X_1, \dots, X_N) du processus des coordonnées. On suppose que pour tout N assez grand, F et G ont des densités F_N, G_N par rapport à la mesure de Lebesgue sur \mathbb{R}^N . Pour qu'il en soit ainsi, il faut et il suffit que F et G aient des densités spectrales f, g telles que $\log f, \log g \in L^1(T)$ (s'il n'en est pas ainsi, le problème statistique étudié ici est trivial).

RAPPORTS DE VRAISEMBLANCE

On pose $L_N(t) = \frac{1}{N} \log \phi_N(t)$

avec $\phi_N(t) = E_F \exp t \log \frac{dG_N}{dF_N}$, $t \in [0,1]$

On désigne par $(T_N f)$, $(T_N g)$ la suite des matrices de Toeplitz associées à f et g .

On a $\log \frac{dG_N}{dF_N} = \frac{1}{2} \log \det T_N^{-1} g T_N f + \frac{1}{2} t X^N (T_N^{-1} f - T_N^{-1} g) X^N$

Comme la covariance de $T_N^{-1/2} f X^N$ est I_N (identité sur \mathbb{R}^N), on a en désignant par $(\mu_i^N)_{i=1 \dots N}$ les valeurs propres de $A_N = I_N - T_N^{1/2} f T_N^{-1} g T_N^{1/2} f$,

$$N L_N(t) = \frac{t}{2} \log \det T_N^{-1} g T_N f - \frac{1}{2} \sum_{i=1}^N \log (1 - \mu_i^N t).$$

Pour $|t| < \frac{1}{\sup_{1 \leq i \leq N} |\mu_i^N|}$, (N fixé), on a

$$\begin{aligned} \sum_{i=1}^N \log (1 - \mu_i^N t) &= - \sum_{p=1}^{\infty} \frac{t^p}{p} \text{Tr } A_N^p \\ &= - \sum_{\ell=1}^{\infty} (-1)^\ell \frac{1}{\ell} \text{Tr} (T_N f T_N^{-1} g)^\ell \left(\frac{t}{1-t} \right)^\ell + N \log (1-t) \end{aligned}$$

Soit $z \in \mathbb{C}$. Comme $T_N f$, $T_N g$ sont régulières on a, dans un voisinage ouvert de 0, $T_N (zf + g)$ régulière, et donc [1],

$$\frac{d}{dz} \log \det T_N (zf + g) = \text{Tr } T_N f T_N^{-1} (zf + g)$$

$$\frac{d}{dz^\ell} \log \det T_N (zf + g) \Big|_{z=0} = (-1)^{\ell-1} (\ell-1)! \text{Tr} (T_N f T_N^{-1} g)^\ell$$

d'où

$$\sum_{i=1}^N \log (1 - \mu_i^N t) = - \sum_{\ell=1}^{\infty} \frac{1}{\ell!} \log \det T_N (zf + g) \Big|_{z=0}^{(\ell)} \cdot \left(\frac{t}{1-t} \right)^\ell + N \log (1-t)$$

De l'analyticité au voisinage de 0 de la fonction $z \rightarrow \log \det T_N (zf + g)$, on tire

$$\sum_{i=1}^N \log (1 - \mu_i^N t) = \log \det T_N \left(\frac{t}{1-t} f + g \right) - \log \det T_N g + N \log (1-t)$$

EXPOSÉ 2

d'où

$$N L_N(t) = \frac{t}{2} \log \det T_N^{-1} g T_N f - \frac{1}{2} \left[\log \det T_N \left(\frac{t}{1-t} f + g \right) - \log \det T_N g + N \log(1-t) \right]$$

$$N L_N(t) = \frac{1}{2} \left[t \log \det T_N f + (1-t) \log \det T_N g - \log \det T_N (tf + (1-t)g) \right] \star$$

Cette formule est valable au voisinage de 0.

Mais la fonction $z \rightarrow \log \det (z T_N f + (1-z) T_N g)$ est analytique dans un \mathbb{C} -voisinage de tout point de $]0,1[$, car $\log(tf + (1-t)g) \in L^1$ (par convexité) et donc $T_N(tf + (1-t)g)$ est régulière pour $t \in [0,1]$.

Par suite, par prolongement analytique, la formule est valable pour $t \in]0,1[$ et par continuité pour $t \in [0,1]$. Appliquons maintenant le théorème élémentaire de Szegö [1]: si $\log f \in L^1$, alors

$$\frac{1}{N} \log \det T_N f \rightarrow \frac{1}{2\pi} \int_T \log f(\theta) d\theta, \text{ quand } N \rightarrow +\infty,$$

et donc :

Théorème : Si $\log f, \log g \in L^1(T)$ et $t \in [0,1]$

$$L_N(t) \rightarrow L(t) \quad (N \rightarrow \infty) \quad \text{avec}$$

$$L(t) = \frac{1}{4\pi} \int_{-\pi}^{\pi} [t \log f(\theta) + (1-t) \log g(\theta) - \log (tf(\theta) + (1-t)g(\theta))] d\theta$$

On en déduit, comme dans [0], la formule de Chernoff et les propriétés asymptotiques du lemme de Neyman-Pearson, moyennant la remarque faite en [0] sur le cas où $M_1=0$ et la deuxième remarque ci-après.

Remarque: 1) Si $f, g \in B$, avec $fg > 0$ et B algèbre de Banach définie dans [2], on a plus précisément

$$L_N(t) = L(t) + O\left(\frac{1}{N}\right).$$

Le même type de raisonnement que celui fait en [2] permet de préciser la convergence quand f, g sont p -fois dérivables.

RAPPORTS DE VRAISEMBLANCE

2) Sous ces hypothèses la formule de Chernoff démontrée au chapitre précédent vaut encore H3) remplaçant le fait que $\phi_N(t)$ est définie sur un intervalle contenant 0 en son intérieur (la modification de démonstration est immédiate). On obtient donc les propriétés asymptotiques du lemme de Neyman-Pearson comme indiqué au § I.

Bibliographie :

- [1] GRENANDER, SZEGÖ
Toeplitz form and their applications
Univ. of California press (1958)
- [2] D. DACUNHA-CASTELLE
Remarque sur l'étude asymptotique du rapport de vraisemblance
de deux processus gaussiens stationnaires
Note aux C.R. Acad. Sc. Paris, Série A, février 1979 (à paraître)
- [3] J. COURSOL, D. DACUNHA-CASTELLE
Sur la formule de Chernoff pour deux processus gaussiens
stationnaires,
Note aux C.R. Acad. Sc. Paris, Série A, mars 1979 (à paraître).

Didier DACUNHA-CASTELLE
Département Mathématique - Bât. 425
E.R.A. CNRS 532 "Statistique Appliquée"
Université Paris-Sud (Centre d'Orsay)
91405 ORSAY.