REVUE DE STATISTIQUE APPLIQUÉE

PIERRE HAMMAD

Mesure d'ordre α de l'information au sens de Fisher

Revue de statistique appliquée, tome 26, n° 1 (1978), p. 73-84

<http://www.numdam.org/item?id=RSA_1978__26_1_73_0>

MESURE D'ORDRE α DE L'INFORMATION AU SENS DE FISHER

Pierre HAMMAD
Maître de Conférences
Laboratoire de Mathématiques Appliquées
Université d'Aix-Marseille, 3, av. Robert Schuman, 13621 Aix-en-Provence

RESUME

A l'image des extensions de type α introduites par RENYI pour l'entropie et le gain d'information au sens de SHANNON, on propose, en tentant éventuellement de l'interpréter, une mesure d'ordre α pour l'information (et le "gain" d'information) au sens de FISHER.

INTRODUCTION

On se limite, pour simplifier, au cas de variables aléatoires continues, à une dimension, de densité $f(x, \theta)$ dépendant d'un paramètre θ ($x \in R$, $\theta \in \Theta \subset R$) et pour les extensions de l'information (type α[1], type β[2], ...) à celles de RENYI [3] (mesure d'ordre $\alpha \geqslant 1$). Elles sont définies et notées comme suit :

1) Pour l'entropie I^S de SHANNON et l'information d'ordre α, $I^{S\alpha}$ de RENYI :

$$I^{S\alpha}(f) = \frac{1}{\alpha - 1} \operatorname{Log} \int f^{\alpha}(x, \theta)\, dx \qquad \alpha > 1 \tag{1}$$

$$I^S(f) = \lim_{\alpha \to 1} I^{S\alpha} = \int f \operatorname{Log} f dx \qquad \alpha = 1 \tag{2}$$

2) Pour le gain d'information G^S de SHANNON (résultant du passage de f à ϕ) et son extension d'ordre α, $G^{S\alpha}$:

$$G^{S\alpha}[f \to \phi] = \frac{1}{\alpha - 1} \operatorname{Log} \int \frac{f^{\alpha}}{\phi^{\alpha - 1}}\, dx \qquad \alpha > 1 \tag{3}$$

$$G^S[f \to \phi] = \lim_{\alpha \to 1} G^{S\alpha}[f \to \phi] = \int f \operatorname{Log} \frac{f}{\phi}\, dx \qquad \alpha = 1 \tag{4}$$

Quant à la "distance" D^S entre f et ϕ [4] :

$$D^S[f \leftrightarrow \phi] = G^S[f \to \phi] + G^s[\phi \to f].$$

Mots-clés : **Information d'ordre α. α-échantillons. Gain d'information de Fisher. Distance statistique.**

6

Dans le même ordre d'idées, on propose ici :

1) de définir une mesure d'ordre $\alpha - I_\theta^{F\alpha}(f) -$ extension de l'information $I_\theta^F(f)$ de FISHER (§ 1) ;

2) d'introduire la notion de gain au sens de FISHER $- G_\theta^F[f \to \phi] -$ (et de distance "statistique" $D_\theta^F[f \longleftrightarrow \phi]$) et son extension à l'ordre α notée $G_\theta^{F\alpha}[f \to \phi]$ (§ 2) ;

3) l'interprétation éventuelle de ces fonctions et quelques unes de leurs expressions, dans le cas de lois normales notamment (§ 1 et 2).

Les définitions adoptées pour $I^{F\alpha}$ et $G^{F\alpha}$ sont fondées sur certains résultats [5] portant sur les informations (SHANNON et RENYI) dans les processus markoviens continus [6].

I. INFORMATION D'ORDRE α DE FISHER

1) Définition et propriétés

• Pour une variable aléatoire réelle X admettant une densité de probabilité $f(x, \theta)$, $x \in R$, $\theta \in \Theta \subset R$, la quantité d'information de FISHER fournie par X sur le paramètre θ est donnée par :

$$I_\theta^F(f) = \int_{\Omega_\theta} f(x, \theta) \left(\frac{\partial}{\partial\theta} \operatorname{Log} f(x, \theta)\right)^2 dx \tag{5}$$

$$\Omega_\theta = \{x \in R : f(x, \theta) > 0\},$$

en supposant réalisées les conditions d'existence de l'intégrale.

• Le cas plus intéressant où $\Omega_\theta = \Omega$ ne dépend pas de θ permet, sous des conditions de régularité pour $f(x, \theta)$, d'écrire (5) sous la forme de la moyenne ou de la variance suivantes :

$$I_\theta^F(f) = - E\left[\frac{\partial^2}{\partial\theta^2} \operatorname{Log} f(x, \theta)\right] = \operatorname{Var}\left[\frac{\partial}{\partial\theta} \operatorname{Log} f(x, \theta)\right] ; \tag{6}$$

ce qui permet l'additivité des quantités d'information apportées par des variables aléatoires indépendantes en probabilité.

Définition

On appellera mesure d'ordre α de l'information de FISHER sur le paramètre θ de la loi de densité $f(x, \theta)$, la quantité lorsqu'elle existe :

$$\left.\begin{array}{ll} I_\theta^{F\alpha}(f) = \alpha^2 \left[\int_{\Omega_\theta} f^\alpha dx\right]^{-1} & \int_{\Omega_\theta} f^\alpha \left(\frac{\partial}{\partial\theta} \operatorname{Log} f\right)^2 dx \\ \alpha \in R^+ & \alpha \geqslant 1 \end{array}\right\} \tag{7}$$

• Par analogie avec la définition adoptée de $I^{S\alpha}$, on se limite au cas $\alpha \geqslant 1$, (7) pouvant cependant être étendue au cas $0 < \alpha < 1$. Il résulte de (7) que :

$$\lim_{\alpha \to 1} I_\theta^{F\alpha}(f) = I_\theta^F(f) = \int_{\Omega_\theta} f \left(\frac{\partial}{\partial \theta} \operatorname{Log} f \right)^2 dx. \tag{8}$$

Changement de variable

Si on passe du système de coordonnées x à un système de coordonnées v (transformation continue et biunivoque), l'information $I^{F\alpha}(f)_x$ se transforme en $I^{F\alpha}(g)_v$; si J (x/v) désigne le jacobien de la transformation et donc :

$$dx = J(x/v)\, dv$$

$$g(v, \theta) = |J(x/v)|\, f(x, \theta),$$

on voit aisément que :

$$I^{F\alpha}(g)_v = \alpha^2 \left[\int_{\Omega_\theta} J^\alpha(x/v)\, f^\alpha J(v/x)\, dx \right]^{-1} \int_{\Omega_\theta} J^{\alpha-1}(x/v)\, f^\alpha \left[\frac{\partial}{\partial \theta} \operatorname{Log}(J(x/v)\, f) \right]^2 dx$$

et par suite :

$$I^{F\alpha}(g)_v = \alpha^2 \left[\int_{\Omega_\theta} \frac{f^\alpha}{|J(v/x)|^{\alpha-1}}\, dx \right]^{-1} \int_{\Omega_\theta} \frac{f^\alpha}{|J(v/x)|^{\alpha-1}} \left(\frac{\partial}{\partial \theta} \operatorname{Log} \frac{f}{|J(v/x)|} \right)^2 dx. \tag{9}$$

En particulier :

$$I^F(g)_v = \int_{\Omega_\theta} f \left(\frac{\partial}{\partial \theta} \operatorname{Log} \frac{f}{|J(v/x)|} \right)^2 dx. \tag{10}$$

• L'information de FISHER, si elle est toujours non négative contrairement à l'information de SHANNON donnée en (2), présente cependant le même "inconvénient" : celui de ne pas rester invariante lors d'une transformation des coordonnées (sauf bien entendu dans le cas de transformations particulières évidentes). C'est une des raisons de l'introduction d'un gain d'information au second paragraphe.

2) Interprétation statistique

• Considérons la famille des lois de densité $f(x, \theta)$ telle que l'on puisse mettre $f^\alpha(x, \theta)$ sous la forme :

$$f^\alpha(x, \theta) = h_\alpha(x, \theta)/g_\alpha(\theta) \tag{11}$$

avec :

avec :

$$\left.\begin{array}{ll} g_\alpha(\theta) > 0 & \\ h_\alpha(x, \theta) \geqslant 0 & \forall\alpha \geqslant 1 \\ \displaystyle\int_{\Omega_\theta} h_\alpha(x, \theta)\, dx = 1. & \end{array}\right\} \qquad (12)$$

On définit ainsi la famille des lois de probabilité de densité :

$$\left.\begin{array}{ll} h_\alpha(x, \theta) = f^\alpha(x, \theta)\, g_\alpha(\theta) & \alpha > 1 \\ h_1(x, \theta) = f(x, \theta)\, ; g_1(\theta) = 1 & \alpha = 1. \end{array}\right\} \qquad (13)$$

(7), (11) et (12) entraînent alors que :

$$\left.\begin{array}{l} I_\theta^{F\alpha}(f) = \displaystyle\int_{\Omega_\theta} h_\alpha \left(\frac{\partial}{\partial\theta} \operatorname{Log} \frac{h_\alpha}{g_\alpha}\right)^2 dx \\ \displaystyle\lim_{\alpha\to 1} I_\theta^{F\alpha}(f) = I_\theta^F(f). \end{array}\right\} \qquad (14)$$

La mesure d'ordre α apparaît ainsi comme une information de FISHER (d'ordre 1) résultant, en vertu de (10), d'une transformation de coordonnées de jacobien $g_\alpha(\theta)$.

• En développant le second membre de (14), on obtient :

$$I_\theta^{F\alpha}(f) = \int_{\Omega_\theta} h_\alpha \left(\frac{\partial}{\partial\theta} \operatorname{Log} h_\alpha\right)^2 dx + \int_{\Omega_\theta} h_\alpha \left(\frac{\partial}{\partial\theta} \operatorname{Log} g_\alpha\right)^2 dx$$

$$- 2 \frac{\partial}{\partial\theta} \operatorname{Log} g_\alpha \int_{\Omega_\theta} h_\alpha \frac{\partial}{\partial\theta} \operatorname{Log} h_\alpha\, dx \, ; \qquad (15)$$

soit encore, si l'on se place dans le cas où $\Omega_\theta = \Omega$ ne dépend pas du paramètre θ :

$$I_\theta^{F\alpha}(f) = I_\theta^F(h_\alpha) + \left(\frac{\partial}{\partial\theta} \operatorname{Log} g_\alpha\right)^2, \qquad (16)$$

d'où :

$$I_\theta^F(h_\alpha) \leqslant I_\theta^{F\alpha}(f) \qquad (17)$$

avec :

$$I_\theta^F(h_\alpha) = I_\theta^{F\alpha}(f) \quad \text{si et seulement si} \quad \frac{\partial g_\alpha}{\partial\theta} = 0. \qquad (18)$$

Cas où $\alpha \in \mathbb{N}^*$

Si l'on se limite au cas de mesures d'ordre entier >0, alors, en vertu des relations (16), (17) et (18), on peut énoncer :

1) L'information d'ordre α, $I_\theta^{F\alpha}(f)$, peut être regardée comme l'information fournie sur le paramètre θ par un α-échantillon $(X_1, \ldots, X_\alpha)$ de la variable X de densité f.

2) $I_\theta^F(h_\alpha)$ représente l'information concernant θ fournie par une statistique T de l'α-échantillon $(X_1 \ldots X_\alpha)$ et admettant pour densité de probabilité la fonction $h_\alpha(t, \theta)$.

3) L'information sur θ fournie par la statistique T est inférieure ou égale à celle fournie par l'α-échantillon.

4) Si $\frac{\partial g_\alpha}{\partial \theta} = 0$, ces deux informations sont égales : T contient alors toute l'information sur θ apportée par l'α-échantillon ; ceci demeure vrai, en vertu de (15), même lorsque le paramètre θ intervient dans les bornes d'intégration. On retrouve dans ce cas la définition de l'exhaustivité de la statistique T [7].

Relation avec l'information de RENYI

L'expression de la mesure d'ordre α de RENYI :

$$I^{S\alpha}(f) = \frac{1}{\alpha - 1} \operatorname{Log} \int_{\Omega_\theta} f^\alpha dx \qquad \alpha > 1$$

et la condition de normalisation :

$$\int_{\Omega_\theta} h_\alpha dx = 1$$

entraînent, compte tenu de (11) :

$$g_\alpha(\theta) = e^{(1-\alpha) I^{S\alpha}} \tag{19}$$

Notons au passage le rapprochement avec la notion de pseudo-cardinal d'ordre α et de cardinal vrai (d'ordre 1) [8] : la mesure d'ordre α de l'information statistique peut être regardée comme une pseudo-information dont la limite ($\alpha = 1$) représente l'information-vraie-de FISHER.

Avec (19), (16) s'écrit :

$$I_\theta^{F\alpha}(f) = I_\theta^F(h_\alpha) + (\alpha - 1)^2 \left[\frac{\partial}{\partial \theta} I^{S\alpha}\right]^2 ; \tag{20}$$

d'où l'on tire l'inégalité :

$$\frac{1}{(\alpha - 1)^2} \geqslant \frac{\left[\frac{\partial I^{S\alpha}}{\partial \theta}\right]^2}{I_\theta^{F\alpha}(f)} \qquad \alpha > 1 \tag{21}$$

analogue à celle de CRAMER-RAO pour les valeurs entières de α.

3) Cas de lois normales ou uniformes

a) Soit X une variable aléatoire suivant une loi normale $N(m, \sigma)$ et $I_m^{F\alpha}(f)$ l'information d'ordre α qu'elle fournit sur le paramètre m (σ connu). Avec :

$$f(x, \theta) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-m)^2}{2\sigma^2}},$$

les expressions de $h_\alpha(x, \theta)$ et $g_\alpha(\theta)$ sont évidentes et le calcul montre que :

$$I_m^{F\alpha}(f) = \frac{\alpha}{\sigma^2} = \alpha I_m^F(f). \tag{22}$$

On retrouve, pour $\alpha \in N^*$, l'information sur m contenue dans un α-échantillon de v. a indépendantes et dans ce cas l'additivité des quantités d'information relatives à ces v. a. Il en est ainsi car en fait la relation (18) est vérifiée. Ce n'est plus vrai si on intéresse à l'information $I_\sigma^{F\alpha}(f)$ sur l'écart-type. On obtient dans ce cas, par un calcul simple :

$$I_\sigma^{F\alpha}(f) = \left(1 - \frac{2}{\alpha} + \frac{3}{\alpha^2}\right)\frac{\alpha^2}{\sigma^2} = (\alpha^2 - 2\alpha + 3)\frac{1}{2} I_\sigma^F(f) \quad \alpha \geqslant 1, \tag{23}$$

information qui croît (à partir de $2/\sigma^2$) avec les valeurs de α.

Notons, dans le cas d'une loi normale $N(m, \sigma)$, que

$$I^{S\alpha}(f) = -\operatorname{Log} \sigma\sqrt{2\pi} + \frac{1}{2}\frac{\operatorname{Log} \alpha}{1 - \alpha}$$

et donc $$\left(\frac{\partial I^{S\alpha}}{\partial \sigma}\right)^2 = \frac{1}{2} I_\sigma^F(f) = \frac{1}{\sigma^2},$$

ce qui vérifie en particulier l'inégalité (21).

b) Si X suit la loi uniforme de densité :

$$f(x, \theta) = \frac{1}{\theta} \quad x \in [0, \theta] = E_\theta$$

$$f(x, \theta) = 0 \quad x \notin [0, \theta]$$

on retrouve immédiatement la valeur :

$$I_\theta^{F\alpha}(f) = \frac{\alpha^2}{\theta^2} \tag{24}$$

et la relation :

$$I_\theta^{F\eta}(f) = \eta^2 I_\theta^F(f) \quad \alpha = \eta \in N^*$$

qui met en défaut la propriété d'additivité des informations lorsque l'ensemble E_θ contient le paramètre θ.

II. GAIN D'INFORMATION D'ORDRE α AU SENS DE FISHER

1) Définitions et propriétés

Ce qui précède montre que, lors d'un changement de variable, l'information de FISHER ne reste généralement pas invariante (introduction du jacobien de la transformation) ; d'où, pour y remédier, la nécessité d'introduire dans l'expression (5) une fonction de norme $\phi(x, \theta)$ pour laquelle on conviendra que $\phi(x, \theta)\,dx$ est un invariant. En choisissant pour $\phi\,dx$ une loi élémentaire, on est alors conduit à définir, à l'image des informations de SHANNON et RENYI, les gains suivants :

Gain d'information au sens de FISHER

Soit $f(x,\theta)$, $x \in R$, $\theta \in \Theta \subset R$, la densité de probabilité de la v. a X. Le remplacement de $f(x,\theta)$ par la densité $\phi(x,\theta)$ (supposée pour simplifier, définie sur le même ensemble Ω_θ) fournit sur le paramètre θ un gain d'information donnée par l'expression, lorsqu'elle existe :

$$\left.\begin{aligned} G_\theta^F [f \to \phi] &= \int_{\Omega_\theta} f(x,\theta) \left[\frac{\partial}{\partial \theta} \operatorname{Log} \frac{f(x,\theta)}{\phi(x,\theta)} \right]^2 dx \\ \Omega_\theta &= \{x \in R,\ f(x,\theta) > 0,\ \phi(x,\theta) > 0\}. \end{aligned}\right\} \tag{25}$$

Gain d'information d'ordre α au sens de FISHER

On appellera, d'une façon plus générale, le gain d'ordre α sur θ résultant du passage de $f(x,\theta)$ à $\phi(x,\theta)$, la quantité suivante (en supposant réalisées les conditions d'existence) :

$$G_\theta^{F\alpha} [f \to \phi] = \alpha^2 \left[\int_{\Omega_\theta} \frac{f^\alpha}{\phi^{\alpha-1}} dx \right]^{-1} \int_{\Omega_\theta} \frac{f^\alpha}{\phi^{\alpha-1}} \left(\frac{\partial}{\partial \theta} \operatorname{Log} \frac{f}{\phi} \right)^2 dx \tag{26}$$

$$\alpha \in R, \quad \alpha \geqslant 1.$$

Pour $\alpha = 1$, on retrouve le gain en (25) :

$$\lim_{\alpha \to 1} G^{F\alpha} [f \to \phi] = G^F [f \to \phi]. \tag{27}$$

Le cas particulier où $\phi = \lambda(x)$ ne dépendrait pas de θ redonne évidemment pour $\alpha = 1$ l'information (vraie) de FISHER :

$$\lim_{\alpha = 1} G^{F\alpha} [f \to \lambda(x)] = I_\theta^F (f). \tag{28}$$

Invariance du gain

Notons d'abord que :

$$\forall\, \alpha \geqslant 1 \quad G^{F\alpha} [f \to \phi] \geqslant 0,$$

l'égalité survenant si f est identique à ϕ.

Dans une transformation biunivoque faisant passer de x à v, le gain ne change pas. En effet, si $J(x/v)$ désigne le jacobien de la transformation, il est immédiat, avec

$$\begin{aligned} dv &= J(v/x)\, dx \\ g &= J(x/v)\, f \\ \psi &= J(x/v)\, \phi \end{aligned}$$

que la nouvelle expression du gain est :

$$G^{F\alpha} [g \to \psi]_v = \alpha^2 \left[\int_{\Omega_\theta} \frac{f^\alpha}{\phi^{\alpha-1}} dx \right]^{-1} \int_{\Omega_\theta} \frac{f\alpha}{\phi^{\alpha-1}} \left(\frac{\partial}{\partial \theta} \operatorname{Log} \frac{f}{\phi} \right)^2 dx$$

soit :

$$G^{F\alpha} [g \to \psi]_v = G^{F\alpha} [f \to \phi]_x \tag{29}$$

2) Relations avec les mesures de RENYI

a) Considérons d'abord le cas particulier où $\phi = k(\theta)$ (distribution uniforme sur E_θ) ; en développant (26), on obtient :

$$G_\theta^{F\alpha}[f \to k] = \alpha^2 \left[\int_{\Omega_\theta} f^\alpha \, dx\right]^{-1} \int_{\Omega_\theta} f^\alpha \left(\frac{\partial}{\partial\theta} \operatorname{Log} f\right)^2 dx + \alpha^2 \left(\frac{\partial}{\partial\theta} \operatorname{Log} k\right)^2 - \tag{30}$$
$$- 2\alpha^2 \left(\frac{\partial}{\partial\theta} \operatorname{Log} k\right)\left[\int_{\Omega_\theta} f^\alpha \, dx\right]^{-1} \int_{\Omega_\theta} f^\alpha \frac{\partial}{\partial\theta} \operatorname{Log} f \, dx.$$

Au second membre de (30), le premier terme de la somme n'est autre que l'information $I^{F\alpha}(f)$ définie en (7). Comme en outre dans le dernier terme,

$$- 2\alpha^2 \left[\int_{\Omega_\theta} f^\alpha \, dx\right]^{-1} \int_{\Omega_\theta} f^\alpha \frac{\partial}{\partial\theta} \operatorname{Log} f \, dx = - 2\alpha \frac{\partial}{\partial\theta} \operatorname{Log} \int_{\Omega_\theta} f^\alpha \, dx$$

et que l'information d'ordre α de RENYI est, d'après (1),

$$I^{S\alpha}(f) = \frac{1}{\alpha - 1} \operatorname{Log} \int_{\Omega_\theta} f^\alpha \, dx,$$

(30) s'exprime finalement comme suit :

$$G_\theta^{F\alpha}[f \to k] = I_\theta^{F\alpha}(f) + \alpha \left(\frac{\partial}{\partial\theta} \operatorname{Log} k\right)\left[\alpha \frac{\partial}{\partial\theta} \operatorname{Log} k - 2(\alpha - 1) I^{S\alpha}(f)\right], \tag{31}$$

c'est-à-dire par une fonction des mesures d'ordre α – au sens de FISHER et de RENYI – de l'information.

b) D'une façon générale, posons :

$$\int_{\Omega_\theta} \frac{f^\alpha}{\phi^{\alpha-1}} \, dx = g_\alpha^{-1}(\theta) \qquad g_1(\theta) = 1 \tag{32}$$

en sorte que l'on définit la famille de lois de densité

$$\left.\begin{array}{l} \forall x, \ \forall \theta \ \ h_\alpha(x,\theta) = \dfrac{f^\alpha(x,\theta)}{\phi^{\alpha-1}(x,\theta)} g_\alpha(\theta) \quad \alpha > 1 \\ \displaystyle\int_{\Omega_\theta} h_\alpha(x,\theta) \, dx = 1 \quad h_\alpha \geqslant 0 \qquad g_\alpha > 0 \\ h_1(x,\theta) = f(x,\theta). \end{array}\right\} \tag{33}$$

On a dans ce cas :

$$G_\theta^{F\alpha}[f \to \phi] = \alpha^2 \int_{\Omega_\theta} h_\alpha \left(\frac{\partial}{\partial\theta} \operatorname{Log} \frac{f}{\phi}\right)^2 dx. \tag{34}$$

Comme d'après (33),

$$\forall \ x \ \operatorname{Log} \frac{f}{\phi} = \frac{1}{\alpha} \operatorname{Log} \frac{h_\alpha}{\phi g_\alpha},$$

le gain peut aussi s'écrire :

$$G_\theta^{F\alpha}[f \to \phi] = \int_{\Omega_\theta} h_\alpha \left(\frac{\partial}{\partial\theta} \operatorname{Log} \frac{h_\alpha}{\phi g_\alpha}\right)^2 dx \tag{35}$$

et apparaît ainsi comme un pseudo-gain d'ordre 1, $G_\theta^F [h_\alpha \to \phi g_\alpha]$ (ϕg_α non normée à l'unité). On peut exprimer $G^{F\alpha}$ en fonction d'un gain (vrai) d'ordre 1 et d'un gain d'ordre α de RENYI, à l'image de la relation (20) ; en effet, (35) entraîne :

$$G_\theta^{F\alpha} [f \to \phi] = \int_{\Omega_\theta} h_\alpha \left(\frac{\partial}{\partial \theta} \mathrm{Log} \frac{h_\alpha}{\phi} \right)^2 dx + \left(\frac{\partial}{\partial \theta} \mathrm{Log}\, g_\alpha \right) \left[\frac{\partial}{\partial \theta} \mathrm{Log}\, g_\alpha - 2 \int_{\Omega_\theta} h_\alpha \frac{\partial}{\partial \theta} \mathrm{Log} \frac{h_\alpha}{\phi} dx \right] \tag{36}$$

avec d'après (3)

$$\mathrm{Log}\, g_\alpha(\theta) = (1-\alpha)\, G^{S\alpha} [f \to \phi]. \tag{37}$$

Mais (36) ne présente d'intérêt que si, non seulement $\Omega_\theta = \Omega$ ne dépend pas de θ, mais ϕ ne dépend pas de θ. Dans ce cas :

$$\left. \begin{array}{l} G_\theta^{F\alpha} [f \to \phi] = G_\theta^F [h_\alpha \to \phi] + (\alpha - 1)^2 \left\{ \frac{\partial}{\partial \theta} G^{S\alpha} [f \to \phi] \right\}^2 \\ \phi = \phi(x),\ \Omega_\theta = \Omega,\ \alpha \geqslant 1. \end{array} \right\} \tag{38}$$

3) Distance statistique

La distance ("divergence") entre les lois fdx et ϕdx est définie à partir des gains d'information de SHANNON comme suit :

$$\begin{aligned} D^S [f \leftrightarrow \phi] &= G^S [f \to \phi] + G^S [\phi \to f] \\ &= \int_{\Omega_\theta} (f - \phi)\, \mathrm{Log} \frac{f}{\phi} dx. \end{aligned} \tag{39}$$

On peut de façon analogue introduire une distance au sens statistique relative au paramètre θ, à partir des gains d'information au sens de FISHER :

$$D_\theta^F [f \leftrightarrow \phi] = G_\theta^F [f \to \phi] + G_\theta^F [\phi \to f] \tag{40}$$

$$D_\theta^F [f \leftrightarrow \phi] = \int_{\Omega_\theta} (f + \phi) \left(\frac{\partial}{\partial \theta} \mathrm{Log} \frac{f}{\phi} \right)^2 dx. \tag{41}$$

Comme pour D^S, on a :

$$D^F > 0$$

$$D^F = 0 \quad \text{si } f \equiv \phi$$

$$D^F [f \leftrightarrow \phi] = D^F [\phi \leftrightarrow f],$$

et elle ne satisfait pas forcément à l'inégalité triangulaire. D^F exprime en quelque sorte une différence d'information : celle-ci diminue si les distributions se rapprochent et est nulle lorsqu'elles se confondent.

4) Exemple des lois normales :

Soit :

$$f(x, \theta) = \frac{1}{\sigma_1 \sqrt{2\pi}} e^{-\frac{(x-m)^2}{2\sigma_1^2}}, \quad \phi(x, \theta) = \frac{1}{\sigma_2 \sqrt{2\pi}} e^{-\frac{(x-m)^2}{2\sigma_2^2}},$$

et $G_m^{F\alpha} [f \to \phi]$ le gain d'ordre α sur le paramètre m.

$f^\alpha/\phi^{\alpha-1}$ se décompose aisément, conformément à (33), en un produit d'une fonction :

$$g_\alpha^{-1}(\theta) = \left(\frac{\sigma_2}{\sigma_1}\right)^\alpha \frac{\sigma_1}{\sqrt{|\alpha(\sigma_2^2 - \sigma_1^2) + \sigma_1^2|}}$$

et d'une densité $h_\alpha(x, \theta)$ d'une loi normale $N\left(m, \Sigma = \frac{\sigma_1 \sigma_2}{\sqrt{|\alpha(\sigma_2^2 - \sigma_1^2) + \sigma_1^2|}}\right)$

Le calcul donne alors :

$$G_m^{F\alpha}[f \to \phi] = \alpha^2 \sigma_1^2 \sigma_2^2 \left(\frac{1}{\sigma_1^2} - \frac{1}{\sigma_2^2}\right)^2 \frac{1}{|(1-\alpha)\sigma_1^2 + \alpha\sigma_2^2|} \tag{42}$$

et en particulier ($\alpha = 1$) :

$$G_m^F[f \to \phi] = \sigma_2^1 \left(\frac{1}{\sigma_2^2} - \frac{1}{\sigma_1^2}\right)^2 = \frac{1}{I_m^F(f)} \; [I_m^F(\phi) - I_m^F(f)]^2 \tag{43}$$

où $I_\theta^F(f)$ et $I_\theta^F(\phi)$ sont les quantités d'information de FISHER apportées par f et ϕ sur la moyenne m.

On peut faire le rapprochement de ce gain avec la distance du khi-2 pour les distributions de densité f et ϕ et interpréter G_m^F comme le gain résultant du passage de la distribution empirique à la distribution théorique.

Par symétrie, on a :

$$G_m^F[\phi \to f] = \frac{1}{I_m^F(\phi)} \; [I_m^F(\phi) - I_m^F(f)]^2 \; ; \tag{44}$$

d'où une distance, au sens statistique, sur m :

$$D_m^F[f \leftrightarrow \phi] = [I_m^F(\phi) - I_m^F(f)]^2 \left(\frac{1}{I_m^F(f)} + \frac{1}{I_m^F(\phi)}\right). \tag{45}$$

Notons par ailleurs que le gain d'ordre α donné en (42) se ramène à un gain d'ordre 1 : en vertu de (36) et puisque $\partial/\partial m \operatorname{Log} g_\alpha(m) = 0$, on doit avoir :

$$G_m^{F\alpha}[f \to \phi] = G_m^F[h_\alpha \to \phi]. \tag{46}$$

On le vérifie immédiatement à partir de (42) où intervient la variance de h_α :

$$\Sigma^2 = \frac{1}{\sigma_1^2 \sigma_2^2} [\alpha(\sigma_2^2 - \sigma_1^2) + \sigma_1^2],$$

ce qui entraîne que :

$$G_m^{F\alpha}[f \to \phi] = \Sigma^2 \left(\frac{1}{\Sigma^2} - \frac{1}{\sigma^2}\right)^2,$$

soit l'analogue de (46).

– Supposons enfin que f et ϕ soient respectivement les densités des lois $N(m_1, \sigma)$ et $N(m_2, \sigma)$ et cherchons $G_\sigma^{F\alpha}[f \to \phi]$.

Suivant la décomposition (33), on obtient :

$$g_\alpha(\sigma) = e^{-\frac{\alpha(\alpha-1)}{2\sigma^2}(m_1 - m_2)^2}$$

et
$$h_\alpha(x, \sigma) = N(\alpha m_1 + (1-\alpha) m_2 \,;\sigma).$$

Par suite,

$$G_\sigma^{F\alpha}[f \to \phi] = \frac{\alpha^2}{\sigma^6} \int_{\Omega_\theta} [(x - m_1)^2 - (x - m_2)^2]\, h_\alpha dx \,;$$

ce qui donne finalement :

$$G_\sigma^{F\alpha}[f \to \phi] = \frac{\alpha^2}{\sigma^6}(m_1 - m_2)^2 \,[4\sigma^2 + (2\alpha - 1)^2 (m_1 - m_2)^2] \tag{47}$$

qui croît avec α et l'écart entre m_1 et m_2. On a en outre :

$$G_\sigma^{F\alpha}[f \to \phi] = G_\sigma^{F\alpha}[\phi \to f] \text{ et } D_\sigma^F[f \longleftrightarrow \phi] = 2G_\sigma^F[f \to \phi]$$

avec (48)

$$G_\sigma^F[f \to \phi] = \frac{1}{\sigma^6}(m_1 - m_2)^2 \,[4\sigma^2 + (m_1 - m_2)^2].$$

Naturellement l'interprétation statistique en termes d'α-échantillons ($\alpha \in N^*$) est plus évidente lorsque $I_\theta^{F\alpha}$ ou $G_\theta^{F\alpha}$ peuvent se ramener à une information de FISHER I^F ou un "gain" de FISHER G_θ^F ; autrement dit, lorsque l'information ou le gain de RENYI correspondants ne dépendent pas du paramètre θ : c'est précisément le cas lorsque θ est la moyenne de lois normales $N(\theta, \sigma)$.

BIBLIOGRAPHIE

[1] KAPUR J.N. – Generalised Entropy of Order α and Type β. *The Math. Seminar*, Vol. IV, p. 78-94, 1967.

[2] VARMA R.S. – Generalisation of RENYI's Entropy of order α. *J. of Math. Sciences*, 1, p. 34-48, 1966.

[3] RENYI A. – On the Measures of Information and Entropy. *Proc. Fourth Berkeley Symp. Math.*, Stat. and Prob., 1, p. 547-561, 1961.

[4] KULLBACK S. – Information Theory and Statistics. Dover Publications, New-York, 1968.

[5] HAMMAD P. – Information dans les systèmes à caractère markovien. *Publications économétriques* Vol. IX, fasc. 2, 1976.

[6] HAMMAD P. – Informations de SHANNON ET FISHER pour les processus de diffusion (à paraître). Congrès IFAC on Information and Systems, Compiègne, oct. 1977.

[7] DUGUE D. – Traité de Statistique théorique et appliquée. Masson, Paris, 1958.

[8] VILLE J.A. – Mathématiques économiques et théorie des Systèmes. Chaire d'Econométrie, Université de Paris VI, 1976.