

La loi de Fisher et ses applications aux images de Radar à Synthèse d'Ouverture.

Jean Marie NICOLAS

GET Télécom Paris
46 rue Barrault, 75634 PARIS CEDEX 13, France
nicolas@tsi.enst.fr

Résumé – La distribution de Fisher, qui entre dans la classe des “lois à queue lourde”, est un outil statistique peu utilisé jusqu’à présent car les méthodes traditionnelles d’estimation de paramètres sont souvent inadaptées. Les “statistiques de deuxième espèce” (ou “Log-statistiques”) permettent en revanche d’en estimer aisément les paramètres. Aussi son utilisation pour l’analyse des images de Radar à Synthèse d’Ouverture semble prometteur, tant sur le plan de la modélisation que sur le plan du traitement.

Abstract – The Fisher probability density function, which can be seen as a “heavy tailed” distribution, is seldom used in statistics as classical methods cannot be used to estimate its parameters. Yet, with the help of “Second kind statistics” (also called “log-statistics”), these parameters are easy to estimate. By this way, the Fisher distribution seems to be an interesting tool to model and process SAR images.

1 Introduction

La loi de Fisher $\mathcal{F}[\mu, L, M]$ est une loi bien connue à 3 paramètres[1] :

$$\mathcal{F}[\mu, L, M] = \frac{L}{M\mu} \frac{\Gamma(L+M)}{\Gamma(L)\Gamma(M)} \frac{\left(\frac{Lu}{M\mu}\right)^{L-1}}{\left(1 + \frac{Lu}{M\mu}\right)^{L+M}}$$

μ pouvant être vu comme une valeur moyenne et L et M comme des facteurs de forme (3 exemples de loi de Fisher sont donnés figure 1). On montre aisément que cette loi est solution du système de Pearson (c’est la solution de type 6 [9]). Cependant, son usage est assez limité car ses moments (entiers ou fractionnaires) s’écrivent :

$$m_\nu = \mu^\nu \frac{\Gamma(L+\nu)}{L^\nu \Gamma(L)} \frac{\Gamma(M-\nu)}{M^{-\nu} \Gamma(M)}$$

et ils ne sont donc définis que pour $\nu > -L$ et $\nu < M$. On a donc affaire à une loi à queue lourde et la relation entre moments successifs m_{n+1} et m_n :

$$\frac{m_{n+1}}{m_n} = \mu \frac{(L+n)M}{L(M-n-1)}$$

peut s’avérer impossible à utiliser pour obtenir le jeu de 3 relations nécessaires pour déterminer les paramètres de cette loi. La méthode du maximum de vraisemblance ne peut, elle aussi, être mise en œuvre. Aussi trouve-t-on peu d’applications de cette loi qui, pourtant, présente l’intérêt d’avoir pour limite la loi Gamma quand $M \rightarrow \infty$ et la loi Gamma inverse quand $L \rightarrow \infty$ (on peut démontrer aisément ces propriétés en analysant les cas limites des solutions du système de Pearson).

2 Utilisation des Log-statistiques

Lorsque on étudie une variable aléatoire positive, il peut être intéressant d’utiliser cette propriété (variable définie sur \mathbb{R}^+)

dans un cadre théorique spécifique : celui des “statistiques de deuxième espèce”[6], que l’on peut aussi appeler “Log-statistiques”. On définit ainsi des “fonctions caractéristiques de deuxième espèce” fondées sur l’utilisation et les propriétés de la transformée de Mellin : cette axiomatique permet d’introduire les log-moments et les log-cumulants. Il apparaît alors que la loi Gamma $\mathcal{G}[\mu, L]$:

$$\frac{1}{\Gamma(L)} \frac{L}{\mu} \left(\frac{Lu}{\mu}\right)^{L-1} e^{-\frac{Lu}{\mu}}$$

joue un rôle essentiel dans ce formalisme (analogue à celui que joue la loi normale dans le monde des statistiques habituelles). Notons que ses trois premiers log-cumulants s’écrivent :

$$\begin{aligned} \tilde{\kappa}_{x(1)} &= \log(\mu) + \Psi(L) - \log(L) \\ \tilde{\kappa}_{x(2)} &= \Psi(1, L) \\ \tilde{\kappa}_{x(3)} &= \Psi(2, L) \end{aligned}$$

$\Psi(L)$ désignant la fonction Digamma (dérivée logarithmique de la fonction Gamma) et $\Psi(n, L)$ désignant les fonctions Polygamma (*i.e.* les dérivées de la fonction Digamma).

La convolution de Mellin $\hat{\star}$ (ou convolution multiplicative) est alors la loi de composition qui permet de formaliser les multiplications de variables aléatoires. En particulier, la convolution de Mellin de deux lois Gamma donne la loi K :

$$\mathcal{G}[\mu_L, L] \hat{\star} \mathcal{G}[\mu_M, M] = \mathcal{K}[\mu_L \mu_M, L, M].$$

Ce résultat assez méconnu a été démontré en 1948 par Epstein dans le cas $L = 1, M = 1$ [2], puis généralisé dans [5]. Il permet d’écrire directement les log-cumulants de la loi K comme la somme des log-cumulants de chaque loi Gamma (propriété d’additivité des log-cumulants lors d’une convolution de Mellin).

Il se trouve que dans ce formalisme, la loi de Fisher s’exprime tout simplement comme la convolution de Mellin d’une loi Gamma et d’une loi Gamma inverse [6] :

$$\mathcal{F}[\mu, L, M] = \mathcal{G}[\mu, L] \hat{\star} \mathcal{GI}[1, M] = \mathcal{G}[1, L] \hat{\star} \mathcal{GI}[\mu, M] \quad (1)$$

On en déduit directement les log-cumulants de la loi de Fisher :

$$\tilde{\kappa}_1 = \log \mu + (\Psi(L) - \log L) - (\Psi(M) - \log M) \quad (2)$$

$$\tilde{\kappa}_2 = \Psi(1, L) + \Psi(1, M), \quad (3)$$

$$\tilde{\kappa}_3 = \Psi(2, L) - \Psi(2, M). \quad (4)$$

Il faut noter que les cumulants d'ordre supérieur à 1 ne dépendent que des paramètres de forme L et M . Il est donc possible de déduire les paramètres L et M par une inversion numérique du système (3,4) : cette inversion est aisée à effectuer puisque les fonctions Polygamma sont strictement monotones. Ensuite le facteur μ se déduit de la relation 2. On voit donc que par le biais des statistiques de deuxième espèce, il est possible d'estimer les paramètres d'une loi de Fisher quelles que soient les valeurs des paramètres de forme L et M (rappelons que les log-cumulants de tout ordre existent pour la loi de Fisher [6]).

3 Interprétation et utilisation de la loi de Fisher

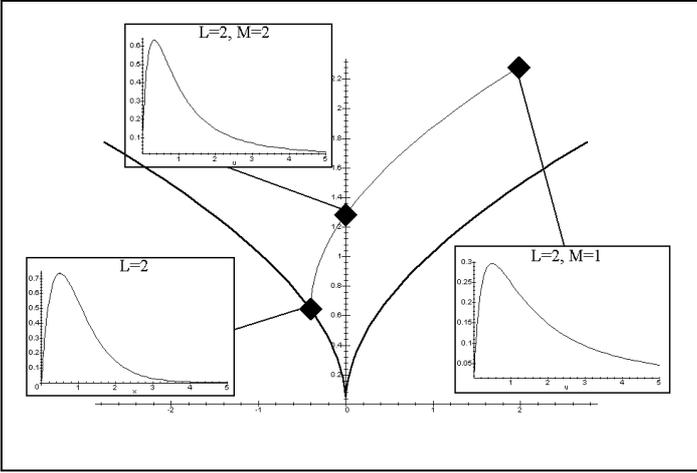


FIG. 1: Diagramme $\tilde{\kappa}_2$ $\tilde{\kappa}_3$ pour la loi de Fisher : cas ($L = 2, M = \infty$), ($L = 2, M = 2$) et ($L = 2, M = 1$). La branche gauche correspond aux lois Gamma, la branche droite aux lois Gamma inverse. Le domaine d'existence des lois de Fisher se trouve entre ces deux branches.

La relation 1 montre que la loi de Fisher se définit à l'aide d'une loi Gamma (qui peut se voir comme une loi "à tête lourde") et d'une loi Gamma inverse (loi à queue lourde). En utilisant la propriété de convergence asymptotique des lois Gamma et Gamma inverse vers la loi homothétique $\mathcal{H}[\mu]$:

$$\lim_{L \rightarrow \infty} \mathcal{G}[\mu, L] = \mathcal{H}[\mu] \quad \lim_{M \rightarrow \infty} \mathcal{GI}[\mu, M] = \mathcal{H}[\mu]$$

on démontre que la loi de Fisher tend vers une loi Gamma si $M \rightarrow \infty$ et vers une loi Gamma inverse si $L \rightarrow \infty$. Une interprétation heuristique de L et M consiste donc à les considérer comme des paramètres de dosage entre un comportement de loi à tête lourde et un comportement de loi à queue lourde.

De manière plus formelle, si on considère le diagramme $\tilde{\kappa}_2$ $\tilde{\kappa}_3$, on voit que la loi de Fisher occupe toute la zone comprise entre la loi Gamma et la loi Gamma inverse. La figure 1 propose plusieurs lois de Fisher (pour $L = 2$) : on voit que le comportement de type "queue lourde" s'accroît lorsque le paramètre M (dosage de loi Gamma inverse) diminue.

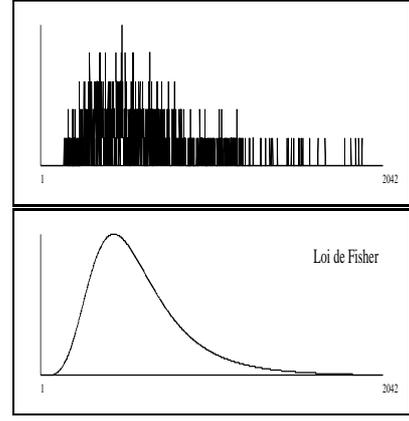


FIG. 2: En haut : histogramme expérimental de données RSO (image ERS sur Paris). En bas : loi de Fisher estimée ($L = 2.69, M = 2.57$).

Aussi, on voit apparaître une raison expérimentale d'utiliser la loi de Fisher. Si les radaristes ont proposé la loi K pour prendre en compte la texture des scènes sous jacentes à une image RSO, c'était pour lever l'hypothèse de stationnarité de la texture nécessaire aux fondements théoriques du chatouement (speckle) modélisé par une loi Gamma. Utiliser la loi de Fisher permet d'une part de couvrir le domaine d'existence de la loi K, mais aussi d'étendre la modélisation aux cas où la loi de probabilité est mal localisée, avec des événements peu fréquents correspondant à de fortes valeurs d'intensité (queue lourde).

Expérimentalement, on voit que cette modélisation semble pertinente : par exemple, sur la figure 2, on voit un histogramme expérimental obtenu sur une image RSO en milieu urbain, et la loi de Fisher estimée à partir de ces échantillons à l'aide des log-cumulants d'ordres 2 et 3, et dont l'aspect semble réaliste. Il est d'ailleurs intéressant de noter que sur cet exemple $L \sim M$, ce qui fait que la loi de Fisher ainsi estimée est en fait très proche d'une loi log-normale (pour laquelle tous les log-cumulants sont nuls à partir de l'ordre 3 : cela correspond donc à l'axe vertical du diagramme $\tilde{\kappa}_2$ $\tilde{\kappa}_3$), choix que propose d'ailleurs Simonetto pour modéliser le milieu urbain [8]. Au cas où la zone d'analyse serait homogène, il faut se rappeler de la convergence asymptotique de la loi de Fisher vers la loi Gamma : le modèle serait alors toujours justifié.

4 Une généralisation de la loi de Rayleigh

En imagerie Radar, le concept de "loi à queue lourde" a été exploité par Kuruoglu et Zerubia [3] qui ont proposé une généralisation de la distribution de Rayleigh. S'appuyant sur des résultats récents concernant les distributions α -stable, les auteurs ont construit cette généralisation sous une forme à deux paramètres (α et γ) dont l'expression ne s'exprime pas sous forme analytique explicite, mais pour laquelle on connaît cependant les moments. Il est alors possible, par prolongement analytique, d'en écrire la fonction caractéristique de deuxième espèce [7] :

$$\phi(s) = \frac{2^s \Gamma\left(\frac{s+1}{2}\right) \gamma^{\frac{s-1}{\alpha}} \Gamma\left(\frac{1-s}{\alpha}\right)}{\Gamma\left(\frac{1-s}{2}\right) \alpha}$$

et d'exprimer les log-cumulants en fonction des deux paramètres α et γ :

$$\begin{aligned}\tilde{\kappa}_1 &= -\Psi(1)\frac{1-\alpha}{\alpha} + \log\left(2\gamma^{\frac{1}{\alpha}}\right) \\ \tilde{\kappa}_p &= \frac{\Psi(p-1, 1)}{\alpha^p} \quad \text{si } p > 1 \text{ pair} \\ \tilde{\kappa}_p &= \frac{\Psi(p-1, 1)}{2^{p-1}} \frac{\alpha^p - 2^{p-1}}{\alpha^p} \quad \text{si } p > 1 \text{ impair}\end{aligned}$$

En particulier, les deux premiers log-cumulants permettent de déduire directement les deux paramètres α et γ sans passer par le schéma itératif proposé dans [3].

Par le biais des log-cumulants, on peut aussi caractériser cette loi pour certaines valeurs de α et nous allons voir le rôle joué par la distribution de Fisher :

– Pour $\alpha = 2$, on retrouve la loi de Rayleigh. En effet :

$$\begin{aligned}\tilde{\kappa}_1 &= \log\left(2\gamma^{\frac{1}{2}}\right) + \frac{\Psi(1)}{2} \\ \tilde{\kappa}_2 &= \frac{\Psi(1, 1)}{4} \\ \tilde{\kappa}_3 &= \frac{\Psi(2, 1)}{8} \dots\end{aligned}$$

C'est donc un cas dégénéré de la loi de Fisher ($M \rightarrow \infty$).

– Pour $\alpha = 1$, les log-cumulants 2 et 3 s'écrivent :

$$\begin{aligned}\tilde{\kappa}_2 &= \frac{\Psi(1, 1)}{4} \\ \tilde{\kappa}_3 &= \frac{-3\Psi(2, 1)}{4}\end{aligned}$$

En utilisant la formule de duplication des fonctions Polygamma, on peut réécrire ces expressions sous la forme suivante :

$$\begin{aligned}\tilde{\kappa}_2 &= \frac{1}{4} \left(\Psi(1, 1) + \Psi\left(1, \frac{1}{2}\right) \right) \\ \tilde{\kappa}_3 &= \frac{1}{8} \left(\Psi(2, 1) - \Psi\left(2, \frac{1}{2}\right) \right)\end{aligned}$$

On reconnaît la formulation (en amplitude) de la convolution de Mellin d'une loi Gamma par une loi Gamma inverse. Par récurrence, on peut étendre la démonstration pour tous les log-cumulants, ce qui est un autre moyen de montrer que cette distribution est la loi de Fisher suivante :

$$\mathcal{F}\left[\mu, 1, \frac{1}{2}\right]$$

– Pour $\alpha = \frac{1}{2}$, considérons le log-cumulant d'ordre 2 :

$$\tilde{\kappa}_2 = 4\Psi(1, 1)$$

On peut réécrire cette expression comme :

$$\tilde{\kappa}_2 = \Psi(1, 1) + \frac{1}{4} \left(\Psi\left(1, \frac{1}{4}\right) + \Psi\left(1, \frac{3}{4}\right) \right) \quad (5)$$

Le log-cumulant d'ordre 3 s'écrit :

$$\tilde{\kappa}_3 = -\frac{31}{4}\Psi(1, 1)$$

On peut réécrire cette expression comme :

$$\tilde{\kappa}_3 = \frac{-3\Psi(2, 1)}{4} - \frac{1}{4} \left(\Psi\left(2, \frac{1}{4}\right) + \Psi\left(2, \frac{3}{4}\right) \right) \quad (6)$$

On reconnaît dans le premier terme des expressions 5 et 6 les log-cumulants du cas $\alpha = 1$, les termes complémentaires correspondant à deux lois Gamma Inverse ($M' = \frac{1}{4}$ et $M'' = \frac{3}{4}$) ce qui donne pour formulation analytique de la loi de Rayleigh généralisée l'expression suivante :

$$\mathcal{F}\left[\mu, 1, \frac{1}{2}\right] \hat{\star} \mathcal{GI}\left[1, \frac{3}{4}\right] \hat{\star} \mathcal{GI}\left[1, \frac{1}{4}\right]$$

– Par récurrence, on peut ainsi montrer que les cas $\alpha = \frac{1}{2^p}$ sont assimilables à des convolutions de Mellin d'une loi de Fisher et de lois Gamma inverse.

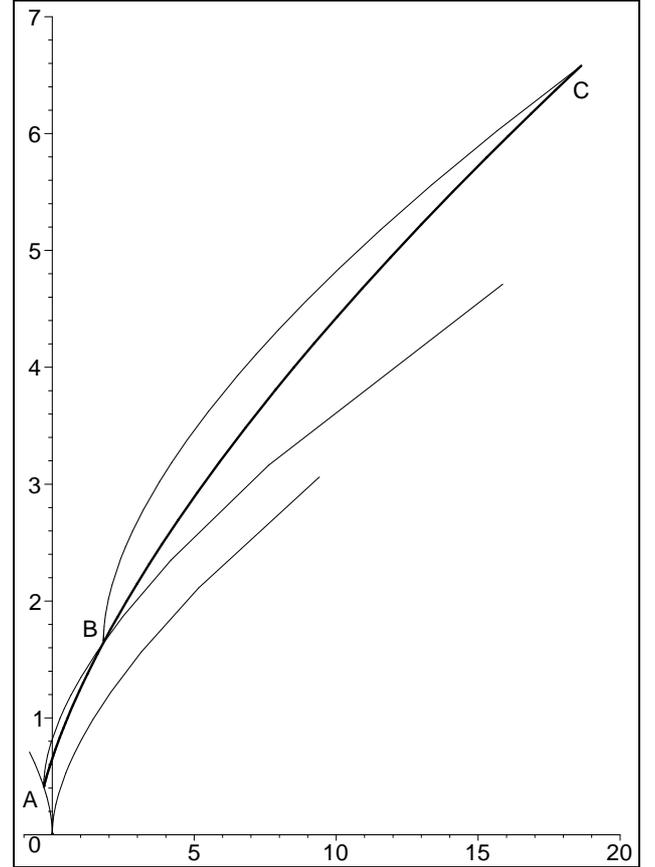


FIG. 3: Diagramme $\tilde{\kappa}_2 \tilde{\kappa}_3$ pour la loi de Rayleigh généralisée en fonction du paramètre α (en gras). Le point A correspond au cas $\alpha = 2$, c'est à dire à une distribution de Rayleigh. Soit la loi de Fisher $\mathcal{F}[\mu, 1, M]$: sa représentation démarre du point A (cas dégénéré $M \rightarrow \infty$) et intersecte en B la loi de Rayleigh généralisée pour la valeur $M = \frac{1}{2}$: ce point correspond à la valeur $\alpha = 1$. Depuis B, on construit la loi $\mathcal{GI}[1, \frac{1}{4}] \hat{\star} \mathcal{GI}[1, M']$: elle intersecte la loi de Rayleigh généralisée (avec $\alpha = \frac{1}{2}$) pour la valeur $M' = \frac{3}{4}$ au point C.

5 Application au filtrage d'images RSO

Puisque la loi de Fisher semble être un modèle bien adapté aux images RSO, il doit être possible de l'introduire dans des techniques de filtrage prenant en compte une modélisation statistique [4]. Parmi celles ci, les filtres MAP (Maximum A Posteriori) peuvent s'avérer faciles d'emploi, mais il est intéressant de noter qu'ils n'ont été appliqué principalement que sur la loi

K (ce qui conduit au filtre γ -MAP), mais jamais sur la loi de Fisher.

Soit un pixel p et les estimées des paramètres d'une loi de Fisher calculée sur un voisinage de ce pixel : $\hat{\mu}$, \hat{L} , \hat{M} . Il est alors aisé de montrer que la valeur \hat{p} maximisant la probabilité a posteriori s'écrit :

$$\hat{p} = \frac{\hat{L}p + \hat{M}\hat{\mu}}{1 + \hat{L} + \hat{M}}$$

ce qui donne au passage un filtre de type adaptatif. Notons que si $M \rightarrow \infty$, $\hat{p} = \hat{\mu}$, ce qui est sensé puisque la texture sous jacente est homogène. Si $M \ll L$, $\hat{p} \rightarrow p$, on attribue une valeur proche de la valeur initiale du pixel, ce qui semble aussi une démarche constructive. Notons enfin que cette expression est beaucoup plus simple que celle obtenue pour le filtre γ -MAP.

Un exemple d'un tel filtre est donné sur une image PRI ERS de Paris (figure 4).

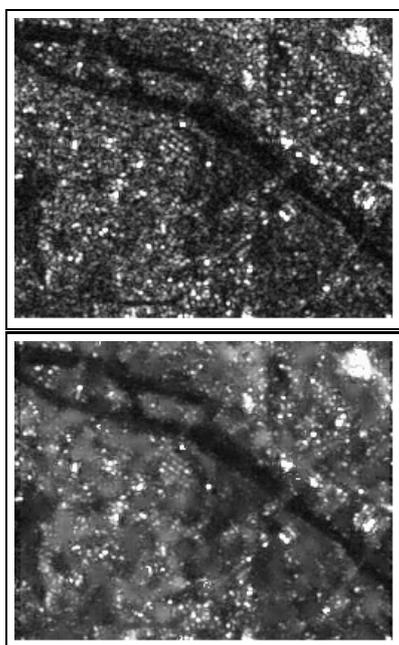


FIG. 4: Image PRI ERS sur Paris (©ESA) et image filtrée par le filtre Fisher-MAP (en bas). L'image a pour dimension 256x192 pixels. La fenêtre d'analyse est de dimension 9x9. La dimension du pixel est de 12,5m x 12,5m.

6 Conclusion et perspectives

Les statistiques de deuxième espèce semblent donner à loi de Fisher de grandes perspectives. En effet elles offrent la possibilité d'estimer les paramètres de cette loi par le biais des log-cumulants, permettant ainsi son utilisation dans un filtre MAP. De plus, la définition de la loi de Fisher comme convolution de Mellin d'une loi Gamma et une loi Gamma inverse propose une démarche heuristique intéressante en l'absence de modèle théorique : c'est donc le cas des images RSO en milieu urbain pour lesquelles on ne peut dégager, aux résolutions actuelles, de modèle de texture crédible et fiable, et pour lesquelles la loi de Fisher permet de combiner les propriétés connues du chatoiement pleinement développé (loi Gamma) avec des propriétés expérimentales de loi à queue lourde.

Références

- [1] H. Cramér *Mathematical methods of statistics* Princeton University Press, 1946
- [2] B. Epstein *Some applications of the Mellin transform in statistics* Annals of Mathematical Statistics, 19 (1948), pp 370-379
- [3] E. Kuruoglu, J.Zerubia *Modelling SAR images with a generalisation of the Rayleigh distribution* INRIA, rapport 4121, Février 2001
- [4] H. Maître *Traitement des images de Radar à Synthèse d'Ouverture* Hermes, 2001
- [5] JM Nicolas, A. Maruani *Speckle well modeled by Mellin transform* PIERS'98, Nantes, juillet 1998
- [6] JM Nicolas *Introduction aux statistiques de deuxième espèce : application des log-moments et des log-cumulants à l'analyse des lois d'images radar* Traitement du Signal, 2002, Vol 19, n3, pp139-167
- [7] J.M. Nicolas, *α -stable positive distributions : a new approach based on second kind statistics*, EUSIPCO2002, Toulouse
- [8] E. Simonetto *Stéréoscopie d'images SAR haute résolution*. Thèse de doctorat, Université de Rennes, 2002
- [9] A.Stuart, J.Keith *Kendall's Advanced Theory of Statistics, Vol.1 Distribution Theory* Griffin, 1987