Introduction to Second Kind Statistics: Application of Log-Moments and Log-Cumulants to the Analysis of Radar Image Distributions

Jean-Marie Nicolas and Stian Normann Anfinsen (translator)

Abstract—Statistical methods classically used to analyse a probability density function (pdf) are founded on the Fourier transform, on which useful tools such the first and second characteristic function are based, yielding the definitions of moments and cumulants. Yet this transformation is badly adapted to the analysis of probability density functions defined on $\mathbb{R}^+$, for which the analytic expressions of the characteristic functions may become hard, or even impossible to formulate. In this article we propose to substitute the Fourier transform with the Mellin transform. It is then possible, inspired by the precedent definitions, to introduce *second kind statistics*: second kind characteristic functions, second kind moments (or log-moments), and second kind cumulants (or log-cumulants). Applied to traditional distributions like the gamma distribution or the Nakagami distribution, this approach gives results that are easier to apply than the classical approach. Moreover, for more complicated distributions, like the $\mathcal{K}$ distributions or the positive α-stable distributions, the second kind statistics give expressions that are truly simple and easy to exploit. The new approach leads to innovative methods for estimating the parameters of distributions defined on $\mathbb{R}^+$. It is possible to compare the estimators obtained with estimators based on maximum likelihood theory and the method of moments. One can thus show that the new methods have variances that are considerably lower than those mentioned, and slightly higher than the Cramér-Rao bound.

Index Terms—Probability density functions defined on $\mathbb{R}^+$, gamma distribution, Nakagami distribution, characteristic functions, parameter estimation, Mellin transform

1 Introduction

Estimation of the parameters of a probability density functions (pdf) is a topic of major significance in pattern recognition. Starting from these estimates, segmentation and classification algorithms can be implemented, both in the field of signal processing and image processing. In signal processing, the intrinsic knowledge of the nature of the data (provided by an acoustic sensor, electromagnetic sensor, etc.) allows us to make realistic assumptions about the suitable distribution models. In particular, many techniques are based on the additive noise model with a noise term that is assumed to be Gaussian. Traditionally, if one describes a random phenomenon by a pdf, one will also introduce the concept of the characteristic function, defined as the Fourier transform $\mathcal{F}$ of the pdf. For example, if $p_x(u)$ is the pdf modelling a random variable X, the characteristic function $\Phi_x(v)$ is obtained by the relation [1]:

$$\Phi_x(v) = \mathcal{F}[p(u)](v) = \int_{-\infty}^{+\infty} e^{jvu} p_x(u)\, du\,. \tag{1}$$

The second characteristic function is defined as the logarithm of the characteristic function:

$$\Psi_x(v) = \log(\Phi_x(v))\,. \tag{2}$$

By taking account of properties of the Fourier transform, it is easy to show that moments of order n are obtained by derivation of the characteristic function:

$$\begin{aligned} m_n &= \int_{-\infty}^{+\infty} u^n p_x(u) du \\ &= (-j)^n \left.\frac{d^n \Phi_x(v)}{dv^n}\right|_{v=0} \end{aligned} \tag{3}$$

and cumulants of order n by derivation of the second characteristic function:

$$\kappa_{x(r)} = (-j)^n \left.\frac{d^r \Psi_x(v)}{dv^r}\right|_{v=0}\,.$$

Moreover, if a phenomenon is analysed, described by a pdf q_y, which is perturbated by an additive noise, described by its pdf r_z, one knows that the output signal is described by the pdf p_x given as

$$p_x = q_y * r_z\,, \tag{4}$$

with the operator $*$ denoting convolution. It is known that the characteristic functions and the cumulants can be written:

$$\Phi_x(s) = \Phi_y(s)\Phi_z(s) \tag{5}$$

$$\Psi_x(s) = \Psi_y(s) + \Psi_z(s) \tag{6}$$

$$\kappa_{x(r)} = \kappa_{y(r)} + \kappa_{z(r)}\ \forall r \tag{7}$$

The author is with the École Nationale Supérieure des Télécommunications, Département TSI, 46 rue Barrault, 75634 Paris cedex 13 (e-mail: jean-marie.nicolas@telecom-paristech.fr).
The original paper was published as: J.-M. Nicolas, "Introduction aux statistiques de deuxième espèce: applications des logs-moments et des logs-cumulants à l'analyse des lois d'images radar", Traitement du Signal, vol. 19, no. 3, pp. 139–167, 2002.
Translated from French by Stian Normann Anfinsen, University of Tromsø, Department of Physics and Technology, NO-9037 Tromsø, Norway (e-mail: stian.normann.anfinsen@uit.no). This version is dated 3 May 2012.

However, in image processing the problems are different. It should be noted first of all that the pixel values are positive or zero (we will not discuss here the analysis of images defined by complex values), and that the noise is often multiplicative. Also, the preceding model must undergo some adaptations to be applicable as it is. One approach often proposed is to perform a logarithmic transformation, which is possible since the pdf is defined on $\mathbb{R}^+$. Several remarks can then be made:

- The analytical calculation of the characteristic functions defined on $\mathbb{R}^+$ is sometimes hard, even impossible for certain distributions, as we will see in section 3.2.
- No complete methodology is proposed for logarithmically transformed data. Calculation of moments on logarithmic scale (that one may call log-moments) is carried out analytically starting from Eq. (3). It requires a change of variable (thus a rewriting of the pdf for this new variable) and is carried out in a specific way for each pdf. This approach requires a good knowledge of integral transforms and of the properties of special functions.
- In traditional statistics, the Gaussian distribution is the reference, which corresponds to the log-normal distribution on a logarithmic scale. However, in many examples, this law does not describe the studied phenomenon well. In particular, the speckle (clutter) observed in images obtained by coherent illumination (e.g., laser, radar, ultrasound) follows, for intensity images, the gamma distribution, which we will study in more detail in this article and which tends asymptotically towards a degenerated Gaussian distribution.

As we will show, a new methodology based on another integral transformation, the Mellin transform [2], makes it possible to perform a more effective analysis of practically important distributions defined on $\mathbb{R}^+$. This methodology, that we propose to call *second kind statistics*, uses the same framework as traditional statistics for the definition of the characteristic functions (simply by replacing the Fourier transform with the Mellin transform in Eq. (1)) and the same construction of the moments and cumulants (by derivation of the characteristic functions). This leads naturally to the definitions of *second kind moments* and *second kind cumulants*. We shall see why we propose to call these new entities *log-moments* and *log-cumulants*. Thanks to this approach, it is possible to analyse in a simpler way distributions with two or three parameters that have traditionally been used for imagery: the gamma distribution, Nakagami distribution, and $\mathcal{K}$ distribution. Then, we will see how to tackle more complex problems like the distributions of the Pearson system, additive mixtures and distributions with heavy tails (i.e., distributions for which the moments cannot be defined starting from a certain order[1]). Finally, we will analyse why the parameter estimators of these distributions based on the log-moments and log-cumulants have a lower variance than those obtained from the traditional moments and cumulants.

The remark can be made that a formalism with such similarity to the existing definitions cannot lead to intrinsically new results. It should be stressed that the essential contribution of this framework is to offer a signal and image processing methodology which proves, for certain applications, considerably easier to use than the traditional approaches. The major goal of this article is to illustrate its simplicity of implementation as well as its flexibility in use.

2 Definition of the Second Kind Characteristic Functions

The objective of this section is to propose a formalism of *second kind statistics* based on the Mellin transform and redefine some elements of traditional statistics, namely the characteristic function yielding moments and cumulants, as outlined in the introduction.

2.1 First Characteristic Function of the Second Kind

Let X be a positive-valued random variable whose pdf, $p_x(u)$, is defined for $u \in \mathbb{R}^+$. The *first characteristic function of the second kind* is defined as the Mellin transform $\mathcal{M}$ of $p_x(u)$:

$$\phi_x(s) = \mathcal{M}[p_x(u)](s) = \int_0^{+\infty} u^{s-1} p_x(u)\, du \qquad (8)$$

provided that this integral converges, which is verified in general only for values of s located inside a strip delimited by two lines parallel to the secondary axis, i.e.

$$s = a + jb, \quad a \in]a1; a2[, \quad b \in \mathbb{R}$$

with a_2 commonly approaching $+\infty$, just as a_1 approaches $-\infty$. As the Mellin transform has an inverse [2], knowing $\phi_x(s)$, one can deduce $p_x(u)$ thanks to the relation:

$$p_x(u) = \frac{1}{2\pi j} \int_{c-j\infty}^{c+j\infty} u^{-s} \phi_x(s)\, ds$$

given that c is confined within the strip where the first characteristic function is defined (i.e, $c \in]a_1; a_2[$). Note that if $p_x(u)$ is a pdf, the second kind characteristic function satisfies the fundamental property:

$$\phi_x(s)|_{s=1} = 1\,.$$

1. Translator's remark: Note that the author uses a strict definition of heavy-tailed distributions. An alternative and more common definition is that heavy-tailed distributions are not exponentially bounded. That is, they have heavier tails than the exponential distribution. Since the context of the discussion is modelling of multilook intensity radar data, it would be natural to replace the exponential distribution with the (generalised) gamma distribution in this criterion.

By analogy, the *second kind moments*, $\tilde{m}_\nu$ ($\nu \in \mathbb{N}$) are defined by the relation:

$$\tilde{m}_\nu = \left. \frac{d^\nu \phi_x(s)}{ds^\nu} \right|_{s=1} . \tag{9}$$

By virtue of a fundamental property of the Mellin transform [Col59]:

$$\mathcal{M}[f(u)(\log u)^\nu](s) = \frac{d^\nu \mathcal{M}[f(u)](s)}{ds^\nu}$$

which is evaluated at $s = 1$, the second kind moments can be written in two different ways:

$$\tilde{m}_\nu = \left. \frac{d^\nu \phi_x(s)}{ds^\nu} \right|_{s=1} \tag{10}$$

$$= \int_0^{+\infty} (\log u)^\nu \, p_x(u) \, du . \tag{11}$$

Eq. (11) suggests that we refer to the second kind moments as *log-moments*, which is adopted for the remainder of the article.

We now introduce the second kind mean or *log-mean* $\tilde{\bar{m}}$. This auxiliary variable is defined by the following relation

$$\log \tilde{\bar{m}} = \tilde{m}_1 \quad \Leftrightarrow \quad \tilde{\bar{m}} = e^{\tilde{m}_1} .$$

Note that this entity, which is in fact the geometric mean, takes its values in $\mathbb{R}^+$ (a suitable scale for the variable u), whereas the log-moments take their values in $\mathbb{R}$ (on logarithmic scale). It is thus possible to compare the mean $\bar{m}$ and the log-mean $\tilde{\bar{m}}$, and the practical importance will be demonstrated for the gamma distribution.

Just as one traditionally defines the central moments, we introduce the definition of the central log-moments of order n, $\tilde{M}_n$:

$$\begin{aligned} \tilde{M}_n &= \int_0^{+\infty} (\log u - \tilde{m}_1)^n \, p_x(u) \, du \\ &= \int_0^{+\infty} \left(\log \frac{u}{\tilde{\bar{m}}} \right)^n \, p_x(u) \, du . \end{aligned} \tag{12}$$

In particular, one readily finds the expression

$$\tilde{M}_2 = \tilde{m}_2 - \tilde{m}_1^2 .$$

Thanks to this formalism, it is possible to obtain an analytical expression for the log-moments by simple derivation of the second kind characteristic function. We will look at the classical interpretation of the Mellin transform.

2.2 A First Interpretation of the Mellin Transform

By comparison of the moment definition in Eq. (3) and the definition of the first characteristic of the second kind in Eq. (8), one can write the generalised moments, m_ν:

$$m_\nu = \phi_x(s)|_{s=\nu+1} = \int_0^{+\infty} u^\nu p_x(u) \, du . \tag{13}$$

For $\nu \in \mathbb{N}$, these are the traditional moments. For $\nu \in \mathbb{R}^+$, we have the fractional moments, that have been used by some authors (like the use of FLOM: Fractional Low Order Moments, in [3]). Provided that the Mellin transform is defined for values of $\nu \in \mathbb{R}^-$, it is justified to use *lower order moments* [4]. Lastly, in addition to moments defined for a value $\nu = a$ (i.e., traditional moments, fractional moments, or lower order moments), one can define moments of complex order with $\nu = a + jb$ for all b, this because the pdf $p_x(u)$ is positive by definition, a property which is trivial to verify.

2.3 Second Kind Cumulants or Log-Cumulants

Still by analogy with classical statistic for scalar real random variables defined on $\mathbb{R}$, the *second characteristic function of the second kind* is defined as the natural logarithm of the first characteristic function of the second kind:

$$\psi_x(s) = \log(\phi_x(s)) . \tag{14}$$

The derivative of the second characteristic function of the second kind, evaluated at $s = 1$, defines *second kind cumulants* of order n:

$$\tilde{\kappa}_{x(n)} = \left. \frac{d^n \psi_x(s)}{ds^n} \right|_{s=1} . \tag{15}$$

Since formally, second kind cumulants are constructed according to the same rules as traditional cumulants, the relations between log-moments and log-cumulants are identical to the relations existing between moments and cumulants. For instance, the three first log-cumulants can be written as:

$$\begin{aligned} \tilde{\kappa}_1 &= \tilde{m}_1 \\ \tilde{\kappa}_2 &= \tilde{m}_2 - \tilde{m}_1^2 \\ \tilde{\kappa}_3 &= \tilde{m}_3 - 3\tilde{m}_1\tilde{m}_2 + 2\tilde{m}_1^3 \end{aligned}$$

As in the case of log-moments, we adopt the name *log-cumulants* for the second kind cumulants.

2.4 Some Properties of Log-Moments and Log-Cumulants

The utilisation of the Mellin transform requires knowledge about some of its specific properties. In particular, let us point out the definition of the Mellin convolution (which is an associative and commutative operation):

$$\begin{aligned} h = f \hat{\star} g \;\Leftrightarrow\; h &= \int_0^\infty f(y) \, g\left(\frac{u}{y}\right) \frac{dy}{y} \\ \Leftrightarrow\; h &= \int_0^\infty g(y) \, f\left(\frac{u}{y}\right) \frac{dy}{y} , \end{aligned} \tag{16}$$

Its fundamental property is similar to the convolution property of the Fourier transform:

$$\mathcal{M}[f \hat{\star} g](s) = \mathcal{M}[f](s) \, \mathcal{M}[g](s) .$$

Note that if f and g are pdfs, then h is also a pdf (i.e., $h(u) \geq 0 \;\forall u \in \mathbb{R}^+$ and $\mathcal{M}[h]|_{s=1} = 1$).

The use of this operator finds an immediate application in the study of multiplicative noise. Let Y and Z

be two independent random variables whose respective pdfs, q_y and r_z, are defined on $\mathbb{R}^+$. Consider a random variable X constructed by a multiplication of of these two variables. It is thus a model of multiplicative noise. It is then shown that the pdf of X, p_x, is obtained as the Mellin convolution of q_y and r_z [5], [6]:

$$p_x = q_y \,\hat{\star}\, r_z\,.$$

The properties deduced in the following are formally identical to those obtained in the case of a traditional convolution (Eqs. (5)-(7)). If ϕ_x is the second kind characteristic function of X, ϕ_y is the second kind characteristic function of Y and ϕ_z is the second kind characteristic function of Z, the following relations are obtained:

$$\begin{aligned} \phi_x(s) &= \phi_y(s)\,\phi_z(s) \\ \psi_x(s) &= \psi_y(s) + \psi_z(s) \\ \tilde{\kappa}_{x(n)} &= \tilde{\kappa}_{y(n)} + \tilde{\kappa}_{z(n)} \quad \forall n \in \mathbb{N} \end{aligned} \tag{17}$$

It is noted in particular that in the case of multiplicative noise, the log-cumulants are additive. This property is not surprising since the common method used to handle multiplicative noise, transformation into logarithmic scale, allows us to treat noise of multiplicative nature like additive noise.

Finally note the following property:

$$u\,(f \,\hat{\star}\, g) = (u\, f)\,\hat{\star}(u\, g)\,.$$

One can also, in a step similar to that of traditional convolution, define the inverse convolution (a non-commutative and non-associative operator). If the ratio

$$\frac{\mathcal{M}[f](s)}{\mathcal{M}[g](s)}$$

is defined in the vicinity of $s = 1$ such that the inverse Mellin transform can be evaluated, the following relation is posed:

$$h = f \,\hat{\star}^{-1}\, g \quad \Leftrightarrow \quad \mathcal{M}[h](s) = \frac{\mathcal{M}[f](s)}{\mathcal{M}[g](s)}\,.$$

With the above notation we have, given that the pdfs p_x, q_y and r_z exist:

$$p_x = q_y \,\hat{\star}^{-1}\, r_z\,,$$

from which we deduce:

$$\begin{aligned} \phi_x(s) &= \frac{\phi_y(s)}{\phi_z(s)} \\ \psi_x(s) &= \psi_y(s) - \psi_z(s) \\ \tilde{\kappa}_{x(n)} &= \tilde{\kappa}_{y(n)} - \tilde{\kappa}_{z(n)} \quad \forall n \in \mathbb{N} \end{aligned} \tag{18}$$

Finally, it can be useful to utilise the *Mellin correlation* (also a non-associative and non-commutative operation), which is defined by the relation:

$$h = f \,\hat{\otimes}\, g \;\Leftrightarrow\; \mathcal{M}[h](s) = \mathcal{M}[f](s)\,\mathcal{M}[g](2-s)\,. \tag{19}$$

A pdf must satisfy $\mathcal{M}[h]|_{s=1} = 1$, to which h complies. Starting from this relation and using the same notation, we can, provided that p_x satisfies

$$p_x = q_y \,\hat{\otimes}\, r_z\,,$$

deduce the following:

$$\begin{aligned} \phi_x(s) &= \frac{\phi_y(s)}{\phi_z(2-s)} \\ \psi_x(s) &= \psi_y(s) - \psi_z(2-s) \\ \tilde{\kappa}_{x(n)} &= \tilde{\kappa}_{y(n)} + (-1)^n \tilde{\kappa}_{z(n)} \quad \forall n \in \mathbb{N} \end{aligned} \tag{20}$$

The following expression can then be shown:

$$h = f \,\hat{\otimes}\, g \;\Leftrightarrow\; h = \int_0^\infty f(uy)\, g(y)\, y\, dy\,. \tag{21}$$

We also note the property:

$$u\,(f \,\hat{\otimes}\, g) = (u\, f) \,\hat{\otimes} \left(\frac{g}{u}\right) \tag{22}$$

In fact, the interpretation of the Mellin correlation is founded on the analysis of the inverse distribution, i.e., the distribution $p_I(u)$ of the random variable $Y = 1/X$, where the random variable X follows the distribution $p(u)$. The relation between these distributions are known to be:

$$p_I(u) = \frac{1}{u^2} p\left(\frac{1}{u}\right)\,.$$

By taking account of a fundamental property of the Mellin transform:

$$\mathcal{M}\left[\frac{1}{u} f\left(\frac{1}{u}\right)\right](s) = \mathcal{M}[f(u)](1-s)\,,$$

it is easily deduced that

$$\mathcal{M}[p_I](s) = \mathcal{M}[p](2-s)\,. \tag{23}$$

It is then seen that the Mellin correlation of a pdf q_y of the random variable Y and a pdf r_z of the random variable Z,

$$q_y \,\hat{\otimes}\, r_z\,,$$

is simply a way to establish the pdf of the random variable Y/Z.

Lastly, as for the traditional characteristic function, it is interesting to note that the second kind characteristic function can be expanded in terms of log-cumulants:

$$\begin{aligned} \psi_x(s) = \tilde{\kappa}_{x(1)}(s-1) &+ \frac{1}{2!}\tilde{\kappa}_{x(2)}(s-1)^2 \\ &+ \frac{1}{3!}\tilde{\kappa}_{x(3)}(s-1)^3 + \ldots \end{aligned}$$

2.5 Theorem of Existence of Log-Moments and Log-Cumulants

We have just seen that the theoretical introduction of the log-moments and log-cumulants does not pose any formal problem. However, the existence of these entities has not been proven, and an interrogation into the requirements for their existence is needed. In this section we will present a theorem of strong conditions, that generally verify the existence of the log-moments and log-cumulants for the distribution usually applied in signal and image processing.

TABLE 1
Properties of the Mellin convolution, the inverse Mellin convolution, and the Mellin correlation of two distributions defined on $\mathbb{R}^+$: p_A and p_B, with second kind characteristic functions ϕ_A and ϕ_B, and log-cumulants $\tilde{\kappa}_{A,n}$ and $\tilde{\kappa}_{B,n}$.

	Characteristic function	Cumulant
$p_A \hat{\star} p_B$	$\phi_A(s)\,\phi_B(s)$	$\tilde{\kappa}_{A,n} + \tilde{\kappa}_{B,n}$
$p_A \hat{\star}^{-1} p_B$	$\frac{\phi_A(s)}{\phi_B(s)}$	$\tilde{\kappa}_{A,n} - \tilde{\kappa}_{B,n}$
$p_A \hat{\otimes} p_B$	$\phi_A(s)\,\phi_B(2-s)$	$\tilde{\kappa}_{A,n} + (-1)^n \tilde{\kappa}_{B,n}$

Let $p(u)$ be a probability distribution defined on $\mathbb{R}^+$, whose second kind characteristic function is $\phi(s)$. This pdf satisfies the relations:

- $p(u) \geq 0 \quad \forall u \geq 0$
- $\int_0^{+\infty} p(u)\,du = \phi(s)|_{s=1}$

Theorem 1: If a pdf has a second kind characteristic function defined on the set $\Omega =]s_A, s_B[$, where $s = 1 \in \Omega$, then all of its log-moments and log-cumulants exist.

Proof: The existence of the log-moments and log-cumulants depends by the convergence of the integral

$$\int_0^{+\infty} (\log u)^n p(u)\,du\,.$$

In order to study this improper integral, we will study its behaviour at 0 and at the limit to infinity.

- **close to infinity:** Let $\alpha \in \Omega$ such that $\alpha > 1$. Thus, $\exists\ \alpha > 1$ such that:

$$\phi(\alpha) = \int_0^{+\infty} u^{\alpha-1} p(u)\,du < \infty$$

which amounts to saying that the moments of $p(u)$ (integer order or fractional) can be calculated for all orders between 1 and α. Assume an integer $n \geq 1$. Two cases then arise:

  - For $\forall\ x > 1$ we have $(\log x)^n < x^{\alpha-1}$. In this case, knowing that $p(u)$ is a pdf and satisfies $p(u) \geq 0$, one can write

$$\lim_{b\to\infty} \int_1^b (\log u)^n p(u)\,du$$
$$\leq \lim_{b\to\infty} \int_1^b u^{\alpha-1} p(u)\,du \leq \phi(\alpha)$$

which demonstrates the convergence of the integral as $x \to \infty$.

  - There exists a constant $c > 1$ such that $(\log c)^n = c^{\alpha-1}$. Then, for $\forall\ x > c$ we have $(\log x)^n \leq x^{\alpha-1}$. By an identical argument as for the previous case, we deduce that

$$\lim_{b\to\infty} \int_c^b (\log u)^n p(u)\,du \leq \phi(\alpha)$$

which demonstrates the convergence of the integral as $x \to \infty$.

- **close to 0:**
  - First of all, consider the particular case where the pdf is bounded. Assume that

$$\exists\ A \quad : \quad \forall u \in [0,1]\,, \quad p(u) \leq A\,,$$

and calculate the limit

$$\lim_{a\to\mathbf{0}} \int_a^1 (\log u)^n p(u)\,du\,.$$

Since the pdf is bounded, we have for $\forall a \in]0,1[$ that

$$\left|\int_a^1 (\log u)^n p(u)\,du\right| \leq \left|\int_a^1 (\log u)^n A\,du\right|$$
$$\leq A \left|\int_a^1 (\log u)^n\,du\right|\,.$$

The following property

$$\left|\lim_{a\to 0} \int_a^1 (\log u)^n\,du\right| = \Gamma(n+1)$$

proves the convergence at 0.

  - In the general case, the study of the convergence starts from the variables change $x \to \frac{1}{x}$, thereafter utilising the convergence property that we have just shown for the case $x \to \infty$.

We deduce that if a probability distribution with bounded values possesses moments (fractional or integer ordered) of order strictly larger than 0 and strictly smaller than 0, then all its log-moments and log-cumulants exist. □

Note that a far more elegant and concise proof, founded on the properties of analytical functions, can be worked out without major problems based on the assumption that $\phi(s)$ is holomorphic [7], and thus differentiable up to all orders at $s = 1$.

2.6 Comparison with Logarithmic Transformation

At this stage, one can wonder what the advantages of this new approach are, and whether a simple transformation into logarithmic scale would lead to the same result. We will show that in order to calculate a characteristic function after logarithmic transformation, one effectively has to calculate the Mellin transform of the original pdf.

We shall consider a random variable x with density defined over real positive numbers. Its pdf, $p_x(u)$, is thus defined for $u \in \mathbb{R}$, and the characteristic function is written:

$$\Phi_x(v) = \int_0^{+\infty} e^{jvu} p_x(u)\,du\,.$$

Then perform a logarithmic transformation. The new random variable y is described the pdf $q_y(w)$, defined for $w \in \mathbb{R}$, with $w = \log u$. This pdf results from p_x with the relation given by

$$q_y(w) = e^w p_x(e^w)\,.$$

Now calculate the characteristic function of the random variable y:

$$\begin{aligned}\Phi_y(v) &= \int_{-\infty}^{+\infty} e^{jvw} q_y(w)\, dw \\ &= \int_{-\infty}^{+\infty} e^{jvw} e^{w} p_x(e^w)\, dw \\ &= \int_{0}^{+\infty} e^{jv\log u} p_x(u)\, du \text{ with } u = e^w \\ &= \int_{0}^{+\infty} u^{jv} p_x(u)\, du\,.\end{aligned} \tag{24}$$

The relation in (24) is recognised as the Mellin transform of $p_x(u)$ at $s = 1 + jv$:

$$\Phi_y(v) = \left.\phi_x(s)\right|_{s=1+jv}\,. \tag{25}$$

This relation shows that if one knows the Mellin transform of a pdf (i.e., its second kind characteristic function), then one knows the ordinary characteristic function in logarithmic scale.

On logarithmic scale, moments and cumulants are deduced by differentiation (simple or logarithmic) of expression (25), which is equivalent to what was defined in Eq. (10). This is another way to justify the terms *log-moments* and *log-cumulants*. We note, however, that the second kind statistics represent a generic method to find log-moments and log-cumulants directly without requiring a variable change (logarithmic transformation) and also without having to calculate the new distribution for the transformed variable.

Moreover, we will see that in the cases generally encountered in signal and image processing, and where the entities are defined on $\mathbb{R}^+$, it is easier to calculate the Mellin transform than the Fourier transform. Thus, our approach simplifies the working of the problem. In addition, when the Mellin transform is known, one automatically obtains:

- the moments, by inserting positive integers for the Mellin transform variable s, and
- the log-moments, by differentiating the Mellin transform with respect to s and evaluating at $s = 1$.

This should be appreciated by any practitioner, since, by applying a single transformation to the distribution, both moments and log-moments are obtained.

2.7 Comparison between Integral Transforms

The use in this context of an ignored transform: the Mellin transform, may surprise, since there exist more common invertible transforms, such as the Laplace transform, that could potentially play an important role in the study of distributions defined on $\mathbb{R}^+$. At this stage, it is important to observe what the relations between the Fourier transform ($\mathcal{F}$), the Laplace transform ($\mathcal{L}$), and the Mellin transform ($\mathcal{M}$) are. When it exists, the Laplace transform of a pdf $p(u)$ is written:

$$\mathcal{L}[p(u)](\sigma) = \int_0^{\infty} e^{-\sigma u} p(u)\, du$$

while the first characteristic function of this pdf is written

$$\mathcal{F}[p(u)](v) = \int_{-\infty}^{+\infty} e^{jvu} p(u)\, du\,.$$

The following relation is immediately deduced:

$$\left.\mathcal{F}[p(u)](v)\right|_{v=-\frac{b}{2\pi}} = \left.\mathcal{L}[p(u)](\sigma)\right|_{\sigma=jb}$$

Because the Laplace transform variable is a complex entity, one may consider that the Laplace transform could allow for an analytical continuation of the characteristic function [8]. However, the intrinsic properties of the Laplace transform are the same as those of the Fourier transform. A logarithmic transformation (in which case, one will have to use the bilateral Laplace transform) will in reality turn the Laplace transform into a Mellin transform:

$$\left.\mathcal{L}[\tilde{p}(u)]\right|_{\sigma=a+jb} = \left.\mathcal{M}[p(u)]\right|_{s=-a-jb} = \left.\phi_x(s)\right|_{s=-a-jb}$$

where $\tilde{p}$ is the pdf on logarithmic scale.

There are such strong relations between these transforms that, most likely, nothing fundamentally new will be found by the use of the Mellin transform. Therefore, it seems that the choice should be dictated by practical considerations. We have seen that the Mellin transform makes it possible to obtain traditional moments and log-moments at the same time, without having to derive the pdf on logarithmic scale. Moreover, the Mellin transform of the experimental distributions commonly used in signal and image processing can be found in tables. This justifies a further look into the use of this rather ignored transform. This is the pragmatic view which motivates the derivations of this article.

3 Fundamental Examples

We will illustrate the new approach by applying it to distributions used to model synthetic aperture radar (SAR) images. These are the gamma and the generalised gamma distribution (intensity images with fully developed speckle), the Rayleigh and the Nakagami distribution (amplitude images with fully developed speckle), and finally the $\mathcal{K}$ distribution (an intensity distribution modelling fully developed speckle modulated multiplicatively by gamma distributed texture). Even if some of the results obtained are trivial, it seems important to be able to carry out comparisons with these simple and well-known cases, in particular in order to handle the problem of estimating the distribution parameters, an aspect which will be looked at in Section 5.

3.1 Gamma and Generalised Gamma Distribution

The two parameter gamma distribution, denoted $\mathcal{G}[\mu, L]$, is a type III solution of the Pearson system [9]. It is defined on $\mathbb{R}^+$ as

$$\mathcal{G}[\mu, L](u) = \frac{1}{\Gamma(L)} \frac{L}{\mu} \left(\frac{Lu}{\mu}\right)^{L-1} e^{-\frac{Lu}{\mu}} \quad \text{with } L > 0 \tag{26}$$

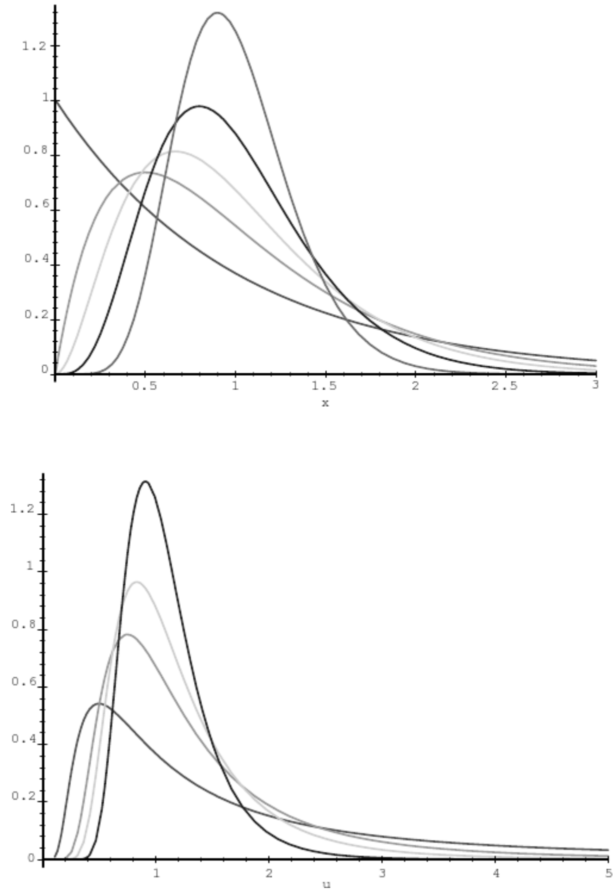


Fig. 1. Top: Gamma distribution $\mathcal{G}[\mu, L]$ with $\mu = 1$ and $L = 1, 2, 3, 5$ and 10. Bottom: Inverse gamma distribution $\mathcal{IG}[\mu, L]$ with $\mu = 1$ and $L = 1, 2, 3, 5$ and 10.

We see that μ is a scale parameter and that L is a shape parameter (Figure 1).

The particular case of $L = 1$ corresponds to the true gamma distribution[2], which is well-known from the radar literature as a model of fully developed speckle in single-look images. The case of $L = \frac{1}{2}$ gives the χ^2 distribution.

The Fourier transform tables show that the characteristic function is written as:

$$\Phi(\nu) = \left(\frac{L}{\mu}\right)^L \frac{e^{jL \arctan\left(\frac{\nu\mu}{L}\right)}}{\left(\nu^2 + \frac{L^2}{\mu^2}\right)^{\frac{L}{2}}} \tag{27}$$

whose complicated expression makes it difficult to use in practice.

On the other hand, by use of known Mellin transforms that can be found in tables [2], [10], the second kind characteristic function can be expressed in terms of the gamma function as:

$$\phi_x(s) = \mu^{s-1} \frac{\Gamma(L + s - 1)}{L^{s-1}\Gamma(L)} \tag{28}$$

2. Translator's remark: Note that the author uses the terms 'generalised gamma distribution' and '(true) gamma distribution' for the distributions more commonly referred to as the 'gamma distribution' and the 'exponential distribution', respectively.

The classical moments $m_n, \forall n \in \mathbb{N}$ are much easier to derive from this function than from (27):

$$m_n = \mu^n \frac{\Gamma(L+n)}{L^n \Gamma(L)} \tag{29}$$

from which we have the well-known moments:

$$m_1 = \mu \qquad m_2 = \frac{L+1}{L}\mu^2 \,.$$

This equation system is analytically invertible, and from the first two moments we derive the following relations for the parameters μ and L:

$$\mu = m_1 \tag{30}$$

$$L = \frac{m_1^2}{m_2 - m_1^2} = \frac{1}{\frac{m_2}{m_1^2} - 1} \tag{31}$$

Note that this distribution is asymmetric, and its mode value is given by:

$$m_{\text{mode}} = \frac{L-1}{L}\mu \leq \mu \,.$$

We also remark that the second kind characteristic function can be separated into a first term, μ^{s-1}, and a second term that only depends on L, the shape parameter. As L goes to infinity, $\frac{\Gamma(L+s-1)}{L^{s-1}\Gamma(L)}$ goes towards 1, and $\mathcal{G}[\mu, L]$ converges in distribution to the homothetic distribution $\mathcal{H}[\mu]$:

$$\begin{aligned} &\phi_x(s) \to \mu^{s-1} \\ \Leftrightarrow\ &\mathcal{G}[\mu, L](u) \to \mathcal{H}[\mu](u) = \frac{1}{\mu}\delta(\mu u - 1) \end{aligned}$$

We note that the homothetic distribution can be seen as a degenerate Gaussian distribution (i.e. with zero variance). It seems to confirm what many experts of radar imaging has pointed out, that the gamma distribution tends towards a Gaussian distribution as L goes to infinity, but by the alternative denotation we avoid abuse of language that can lead to confusion.

Another major point specific to the distributions defined on $\mathbb{R}^+$ rests on the fact that the Mellin transform of $\mathcal{G}[\mu, L]$ is defined for $s > 1 - L$. It is seen that, for $L > 1$, it is possible to have negative values of $s - 1$ and thus *lower order moments*. Qualitatively, the lower order moments – i.e. positive powers of $\frac{1}{u}$ – mainly reflect the weight of the distribution between 0 and μ, while the traditional moments - i.e. positive powers of u - rather analyse the distribution between μ and ∞. Thanks to the lower order moments, it is possible to analyse selectively the left or the right tail of a probability distribution. The importance of this observation for asymmetrical distributions such as the gamma distribution is evident.

The first two log-cumulants of $\mathcal{G}[\mu, L]$ are expressed by the following relations, where $\Phi(\cdot)$ is the digamma function and $\Psi(r, \cdot)$ is the r-th order polygamma function, i.e. the r-th order derivative of the digamma function:

$$\tilde{\kappa}_{x(1)} = \log(\mu) + \Psi(L) - \log(L) \tag{32}$$

$$\tilde{\kappa}_{x(2)} = \Psi(1, L) \geq 0 \tag{33}$$

$$\tilde{\kappa}_{x(3)} = \Psi(2, L) \leq 0 \tag{34}$$

and it is trivial to show that

$$\tilde{\kappa}_{x(r)} = \Psi(r-1, L) \quad \forall r > 1\,,$$

which expresses the property that the log-cumulants depend only on L from second order and upwards.

We note that the property

$$\lim_{L\to\infty} (\Psi(L) - \log(L)) = 0$$

associated with the fact that the polygamma functions go towards 0 at infinity, can be used to show that the gamma distribution converges towards the homothetic distribution as L goes to infinity.

Remark that the third order log-cumulant is negative. This illustrates that, for the gamma distribution, the left tail is heavier than the right tail of the distribution, which decreases very quickly as the argument approaches infinity.

The log-mean is written:

$$\tilde{\tilde{m}} = \mu \frac{e^{\Psi(L)}}{L} \tag{35}$$

It is interesting to note the two following points:

- $\frac{e^{\Psi(L)}}{L} \leq 1$: The log-mean is less than the mean value. Note that this property is valid for all L.
- $\frac{e^{\Psi(L)}}{L} \geq \frac{L-1}{L}$: The log-mean is larger than the mode value.

A more complete analysis would show that the log-mean is also lower than the median value, defined by

$$\int_0^{m_{med}} \mathcal{G}(u)\, du = 0.5\,.$$

It can also be justified to use the log-mean instead of the traditional mean. This gives interesting results in certain applications of image processing [11].

Finally, by a logarithmic transformation, the gamma distribution $\mathcal{G}[\mu, L](u)$ becomes the Fisher-Tippett distribution $\mathcal{G}_{\mathcal{FT}}[\tilde{\mu}, L](w)$ with $\tilde{\mu} = \log \mu$ and $w = \log u$:

$$\mathcal{G}_{\mathcal{FT}}[\bar{\mu}, L](w) = \frac{L^L}{\Gamma(L)} e^{L(w-\bar{\mu})} e^{-Le^{(w-\bar{\mu})}}$$

Its characteristic function is obtained by taking the Fourier transform. Unfortunately, the required relation is not found in tables. This is commonly circumvented by showing that the evaluation amounts to calculating a Mellin transform. In effect, one applies (25) unknowingly.

To conclude, it is seen that in the case of the generalised gamma distribution, the second kind statistics approach allows us:

- to obtain a simpler expression for the second kind characteristic function than for the classical characteristic function.
- to estimate the distribution parameters more efficiently by inversion of Eqs. (32) and (33).
  - The shape parameter L is easily derived from the second order log-cumulant, even if no analytical formulation can be found, since the polygamma functions are monotonous and easy to invert numerically (Tabulation can also be used to save computation time). The variance of this estimator is evaluated in Section 5, and we will see that it is notably lower than the variance obtained with the method of moments estimator, as defined in (31).
  - After L is known, μ can be derived from the expression of the first order log-cumulant.
- to propose a "typical value" for use in image processing, lying between the mode and the mean, which realistically represents a sample if it can be regarded as homogeneous.

3.2 Rayleigh and Nakagami Distribution

We will now handle a problem specific to SAR imagery, namely the transformation of intensity data to amplitude data. Even if models have simple expressions for intensity data (the gamma distribution is known to all scientific communities), the images are quite often available as amplitude data, which will reveal new problems regarding parameter estimation. In this article, we will thus address the transformation from intensity data that follow the gamma distribution, to amplitude data with their resulting distribution.

The Nakagami distribution[3] is the name which in the radar literature has been associated with amplitude data that follow a gamma distribution when transformed into the intensity domain. It is thus a two parameter distribution: $\mathcal{RN}[\mu, L]$, given by:

$$\mathcal{RN}[\mu, L](u) = \frac{2}{\mu} \frac{\sqrt{L}}{\Gamma(L)} \left(\frac{\sqrt{L}u}{\mu}\right)^{2L-1} e^{-\left(\frac{\sqrt{L}u}{\mu}\right)^2}. \tag{36}$$

For $L = 1$, one retrieves the Rayleigh distribution:

$$\mathcal{RN}[\mu, L{=}1](u) = \frac{2}{\mu}\left(\frac{u}{\mu}\right) e^{-\left(\frac{u}{\mu}\right)^2}.$$

The fundamental relation between the Nakagami distribution (for amplitude) and the generalised gamma distribution (for intensity, i.e. squared amplitude) is obtained by a simple variable change, which can be written as:

$$\mathcal{RN}[\mu, L](u) = 2u\mathcal{G}[\mu_{\mathcal{G}}, L](u^2)\,. \tag{37}$$

By means of the following two Mellin transform properties [2]:

$$\mathcal{M}[u^a f(u)](s) = \mathcal{M}[f(u)](s+a)$$
$$\mathcal{M}[f(u^a)](s) = \frac{1}{a}\mathcal{M}[f(u)]\left(\frac{s}{a}\right)$$

3. It is important to return to Nakagami the paternity of this distribution described by two parameters: mean and shape parameter, which has often been wrongly associated with the generalised gamma distribution. The formalism was proposed in 1942 by Nakagami in an exhaustive study of the "m-distribution". It was not published in English until 1960 [12]. Of course, this distribution is for instance found in [13] as the result of transformations starting from the gamma distributions. However, it seems that Nakagami performed the first complete study.

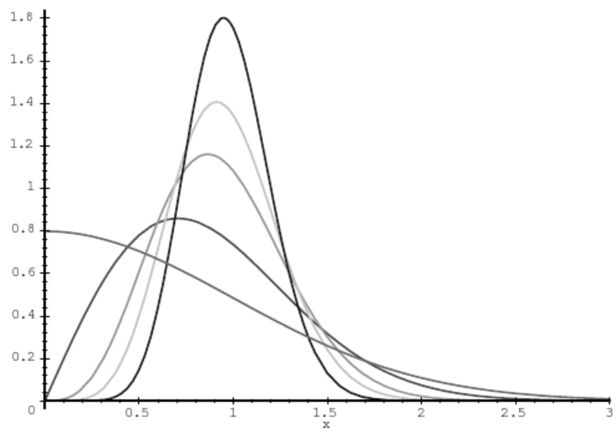


Fig. 2. Rayleigh-Nakagami distribution $\mathcal{RN}[\mu, L]$ with $\mu = 1$ and $L = 0.5, 1, 2, 3, 5$.

and knowing $\phi_{\mathcal{G},x}(s)$, the second kind characteristic function of the gamma distribution, the second kind characteristic function of the Nakagami distribution can be derived directly as:

$$\phi_{\mathcal{RN},x}(s) = \phi_{\mathcal{G},x}\left(\frac{s+1}{2}\right)$$

which, by inserting $\mu_{\mathcal{G}} = \mu^2$, allows us to write:

$$\phi_{\mathcal{RN},x}(s) = \mu^{s-1}\frac{\Gamma(\frac{s-1}{2}+L)}{L^{\frac{s-1}{2}}\Gamma(L)} .$$

This reasoning applies also elsewhere, regardless of the power to which u is raised in the change of variable. It is easily shown that for $v = u^\alpha$, we have

$$p_u(u) = \alpha u^{\alpha-1} p_v(u^\alpha)$$

and the second kind characteristic function of the random variable u is derived directly from the properties of the Mellin transform as:

$$\phi_u(s) = \phi_v\left(\frac{s+\alpha-1}{\alpha}\right) . \tag{38}$$

Note that this result would be useful for the analysis of the Weibull distribution [14], another well-known radar distribution, which we will not address in this article.

The classical moments of the Nakagami distribution follow directly from $\phi_{\mathcal{RN},x}(s)$:

$$m_1 = \frac{\Gamma\left(L+\frac{1}{2}\right)}{\sqrt{L}\Gamma(L)}\mu \qquad m_2 = \mu^2 .$$

Take note of a peculiarity of this distribution: There is a very simple relation between the parameter μ and the second order moment, not the first order moment. On the other hand, the implicit expression of L obtained through the first order moment is very hard to handle. We cannot obtain a simple inversion formula (as in the gamma distribution case, where (31) gave L directly in terms of m_1 and m_2) to solve for L.

The mode of this pdf is

$$m_{\text{mode}} = \sqrt{\frac{2L-1}{2L}}\mu$$

The log-cumulants are derived directly from those of the gamma distribution as:

$$\begin{aligned}
\tilde{\kappa}_{\mathcal{RN},x(r)} &= \left.\frac{d^r\psi_{\mathcal{RN}}(s)}{ds^r}\right|_{s=1} \\
&= \left.\frac{d^r \log\phi_{\mathcal{RN}}(s)}{ds^r}\right|_{s=1} \\
&= \left.\frac{d^r \log\phi_{\mathcal{G}}\left(\frac{s+1}{2}\right)}{ds^r}\right|_{s=1} \\
&= \left(\frac{1}{2}\right)^r \left.\frac{d^r \log\phi_{\mathcal{G}}(s')}{ds'^r}\right|_{s'=1} \\
&= \left(\frac{1}{2}\right)^r \tilde{\kappa}_{\mathcal{G},x(r)}
\end{aligned}$$

From this we deduce:

$$\begin{aligned}
\tilde{\kappa}_{x(1)} &= \log(\mu) + \frac{1}{2}\Psi(L) - \frac{1}{2}\log(L) \\
\tilde{\kappa}_{x(2)} &= \frac{1}{4}\Psi(1, L)
\end{aligned}$$

and for all $r > 1$:

$$\tilde{\kappa}_{x(r)} = \left(\frac{1}{2}\right)^r \Psi(r-1, L)$$

More generally, it is shown for $v = u^\alpha$ that

$$\tilde{\kappa}_{p_u,x(r)} = \left(\frac{1}{\alpha}\right)^r \tilde{\kappa}_{p_v,x(r)}$$

In this case, L can be calculated directly if the second order log-cumulant is known. The problem we meet is of the same kind as for the gamma distribution, namely inversion of polygamma functions.

The log-mean is written

$$\tilde{\tilde{m}} = \mu\frac{e^{\frac{\Psi(L)}{2}}}{\sqrt{L}} \tag{39}$$

To conclude, the motivation of our approach is seen from the fact that the analytical expressions of the log-moments and the log-cumulants have a complexity comparable with the case of the gamma distribution, which is not the case in traditional statistics, where a simple relation between the first two moments and the shape parameter L cannot be obtained.

3.3 Inverse Gamma Distribution

The inverse gamma distribution is another two parameter distribution which is also a solution of the Pearson system (the type V solution). Is is expressed as

$$\mathcal{IG}[\nu, L](u) = \frac{1}{\Gamma(L)}\frac{1}{L\nu}\left(\frac{L\nu}{u}\right)^{L+1} e^{-\frac{L\nu}{u}}$$

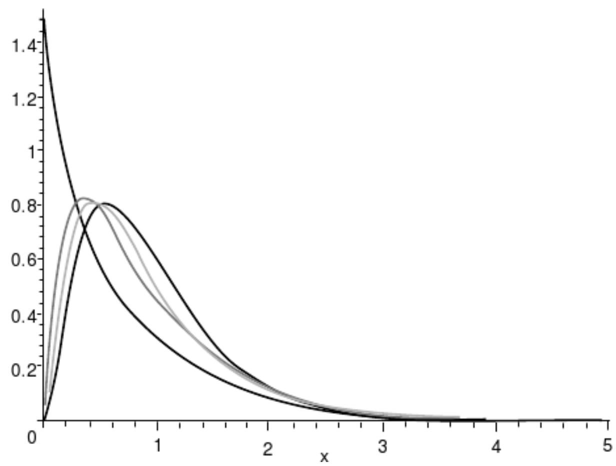


Fig. 3. $\mathcal{K}$ distribution: Fully developed speckle ($L = 3$) modulated multiplicatively by a Rayleigh distribution with $\mu = 1$ and $1, 3, 5$ and 10.

where $L \geq 0$ and $\nu > 0$. Its second kind characteristic function is written:

$$\phi_x(s) = \nu^{s-1} \frac{\Gamma(L+1-s)}{L^{1-s}\Gamma(L)}\,.$$

It is seen that the n-th order moments of the inverse gamma distribution are not defined for $n \geq L$. The inverse gamma distribution is thus an example of heavy tailed distributions. Its log-cumulants, that exist for all orders, are written:

$$\begin{aligned}
\tilde{\kappa}_{x(1)} &= \log(\nu) - \Phi(L) + \log(L)\\
\tilde{\kappa}_{x(2)} &= \Psi(1, L)\\
\tilde{\kappa}_{x(r)} &= (-1)^r \Psi(r-1, L) \quad \forall\, r > 1
\end{aligned}$$

For even r, these are the same as those of the gamma distribution. For odd r, they are opposite (See the more general relation in (20)). As for the gamma distribution, from second order and upwards, the log-cumulants depend only on the shape parameter L. Note that the third order log-cumulant is positive, which is a sufficient condition for being heavy-tailed.

This distribution could also have been introduced as the inverse of the gamma distribution (cf. Section 2.4), which would make it possible to deduce the log-cumulants directly. However, it was important to recall that the inverse gamma distribution is a particular solution of the Pearson system and associated with its own share of work in the literature.

3.4 $\mathcal{K}$ Distribution

With the $\mathcal{K}$ distribution, we will show that second order statistics provide an estimation method for the parameters of a complex distribution by simple application of the results already achieved for the gamma distribution. The $\mathcal{K}$ distribution $\mathcal{K}[\mu, L, M]$ has three parameters and is defined as

$$\begin{aligned}
\mathcal{K}[\mu, L, M](u) = &\frac{1}{\Gamma(L)\Gamma(M)} \frac{2LM}{\mu} \left(\frac{LMu}{\mu}\right)^{\frac{M+L}{2}-1}\\
&\times K_{M-L}\left[2\left(\frac{LMu}{\mu}\right)^{1/2}\right]
\end{aligned} \tag{40}$$

where $K_n(\cdot)$ is the modified Bessel function of the second kind with order n. On this form, calculations of moments and log-moments require good knowledge of Bessel function properties as well as tables of transforms of Bessel functions.

In fact, the $\mathcal{K}$ distribution is the distribution which is followed by a random variable defined as the product of two independent variables that are both gamma distributed. Note that this definition made it possible for Lomnicki [15] to retrieve Eq. (40) using, already at this time, the Mellin transform.

More precisely, it is possible to define the $\mathcal{K}[\mu, L, M]$ distribution as a Mellin convolution of two gamma distributions [6]:

$$\mathcal{K}[\mu, L, M] = \mathcal{G}[1, L] \,\hat{\star}\, \mathcal{G}[\mu, M]$$

This definition greatly simplifies the calculations of the second kind characteristic function and thus the moments and log-cumulants. In effect, from the properties of the Mellin convolution (Section 2.4) and knowing the characteristics of the gamma distribution, one can write the second kind characteristic function of the $\mathcal{K}$ distribution like a product of the second kind characteristic functions of the gamma distributions $\mathcal{G}[1, L]$ and $\mathcal{G}[\mu, M]$:

$$\phi_x(s) = \mu^{s-1} \frac{\Gamma(L+s-1)}{L^{s-1}\Gamma(L)} \frac{\Gamma(M+s-1)}{M^{s-1}\Gamma(L)} \tag{41}$$

which allows us to immediately deduce the classical moments m_1 and m_2 without using the definition of the $\mathcal{K}$ distribution, and thus without needing to know the properties of the Bessel function:

$$m_1 = \mu \qquad m_2 = \mu^2 \frac{L+1}{L} \frac{M+1}{M}\,.$$

In the same manner, we obtain directly the first two log-cumulants as the sum of the log-cumulants of the gamma distributions $\mathcal{G}[1, L]$ and $\mathcal{G}[\mu, M]$:

$$\tilde{\kappa}_{x(1)} = \log\mu + (\Psi(L) - \log(L)) + (\Psi(M) - \log M) \tag{42}$$
$$\tilde{\kappa}_{x(2)} = \Psi(1, L) + \Psi(1, M) \tag{43}$$
$$\tilde{\kappa}_{x(3)} = \Psi(2, L) + \Psi(2, M) \tag{44}$$

and we can show that for all $r > 1$:

$$\tilde{\kappa}_{x(r)} = \Psi(r-1, L) + \Psi(r-1, M)\,.$$

Finally, the normalised second order moment is written:

$$\tilde{M}_2 = \Psi(1, L) + \Psi(1, M) \tag{45}$$

and the log-mean:

$$\tilde{\tilde{m}} = \mu \frac{e^{\Psi(L)}}{L} \frac{e^{\Psi(M)}}{M}\,. \tag{46}$$

Also here, it is easy to derive the shape parameters L and M by virtue of the second and third order log-cumulants (Eqs. (43) and (44)). A simple numerical method is proposed and tested in [14]. The scale parameter is derived from the first order log-cumulant (Eq. (42)).

This method is much easier than the traditional method of moments, which results in a third degree equation. Note also that maximum likelihood estimation cannot be applied for this distribution [16].

4 Applications

Second kind statistics prove easy to put into practice in the framework of fundamental probability distribution defined on $\mathbb{R}^+$. Except for the introduction of the gamma function and its logarithmic derivatives (the polygamma functions), the obtained expressions contain no difficult terms. On the contrary, they are simple and easy to comprehend.

We will now look at several innovative applications of this model:

- A new approach to analysis of the three parameter distributions used to model SAR imagery
- The case of additive mixtures of gamma distributions, for which the traditional approaches lead to expressions that are very hard to handle
- The case of positive α-stable distributions, used for instance by Pierce to characterise sea clutter [17], for which it is difficult to estimate the parameters. (The analytical expression of the pdf for such heavy-tailed distributions is generally not known.)
- Finally, another example of the α-stable distribution proposed by Kuruoğlu and Zerubia [18], which can be seen as a generalisation of the Rayleigh distribution.

We will start by pointing out a method classically used to characterise these distributions: the use of the parameters of asymmetry β_1 (skewness) and peakedness β_2 (kurtosis).

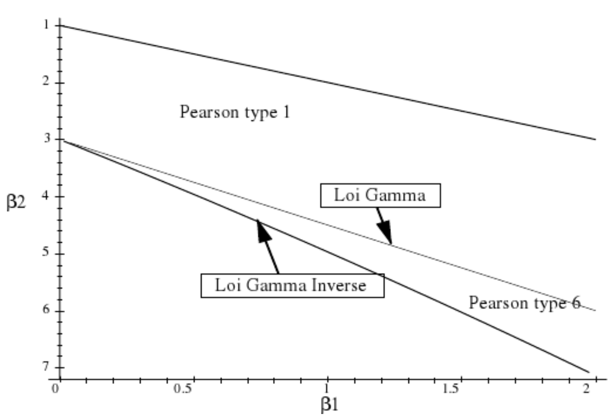


Fig. 4. The Pearson system displayed in a (β_1, β_2) diagram.

4.1 (β_1, β_2) Diagram

Traditionally, the skewness and kurtosis are used to characterise distributions belonging to the Pearson system. These two coefficients are written in terms of the second, third and fourth order moment:

$$\beta_1 = \frac{M_3^2}{M_2^3}$$
$$\beta_2 = \frac{M_4}{M_2^2}$$

The curves obtained for the Pearson system are shown in Figure 4 in their classical representation. The characteristic point of ($\beta_1 = 0$, $\beta_2 = 3$) corresponds to the Gaussian case (It is invariant with respect to variance).

Because of the choice of squaring the third order central moment in β_1, this coefficient is not able to distinguish between distributions that have skewness of the same magnitude but with opposite sign. Thus, it is not possible to separate between "standard" distributions and heavy-tailed distributions. Hence, this diagram seems to be badly adapted to the distributions defined on $\mathbb{R}^+$.

4.2 Characterisation of Texture Distributions in the ($\tilde{\kappa}_3\tilde{\kappa}_2$) Diagram

The (β_1, β_2) diagram is founded on the calculation of traditional centred moments and aims at comparing distributions against the reference Gaussian distribution, for which the skewness is zero and the kurtosis is directly related to the variance (σ). It is then natural to propose a similar approach, founded on the functions of second kind statistics. We propose in this section a new method: the (κ_3, κ_2) diagram, that is, the representation of third order log-cumulants versus second order log-cumulants (that are always positive or zero for pdfs defined on $\mathbb{R}^+$).

In this diagram, the origin corresponds to the homothetic distribution. Because of the asymptotic behaviour of the gamma distribution and the inverse gamma distribution at $L \to \infty$, these distributions are represented by curves joining at the origin. As noted, the gamma

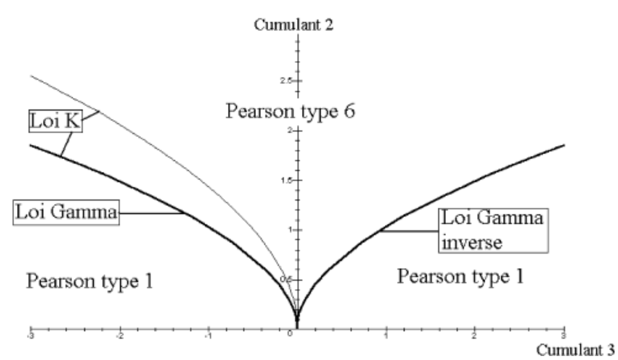


Fig. 5. The Pearson system and the $\mathcal{K}$ distribution displayed in a (κ_3, κ_2) diagram.

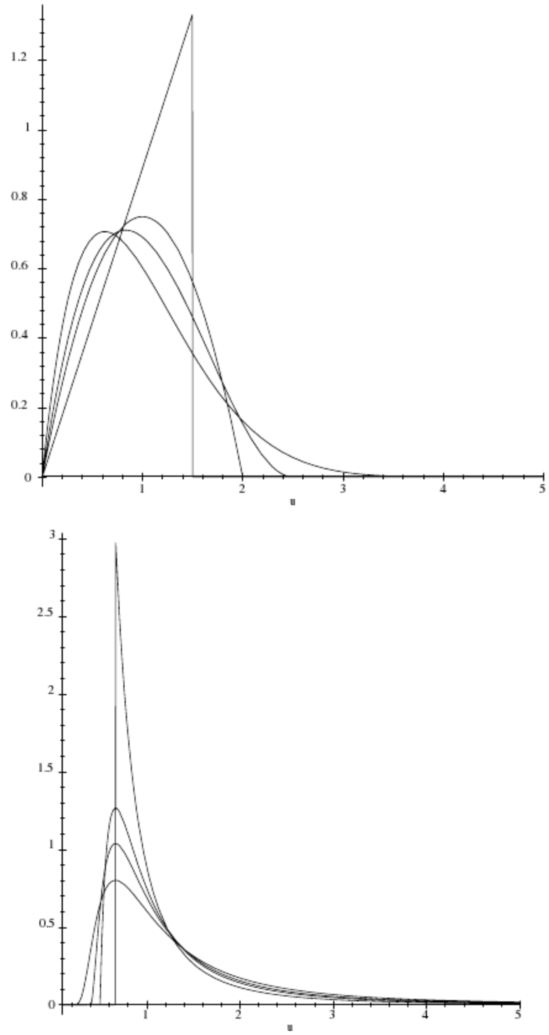


Fig. 6. The Pearson type I distribution (top) and the inverse Pearson type I distribution (bottom).

distribution has negative values for the third-order log-cumulants, while the heavy-tailed inverse gamma distribution has positive values. Note that it is easy to show that the log-normal distribution, for which all log-cumulants of order $n > 2$ is zero, occupies the second axis.

Figure 5 places the gamma distribution and the inverse gamma distribution in the diagram, together with the $\mathcal{K}$ distribution (which occupies a surface above the gamma distribution, limited upwards by a curve defined by the distribution $\mathcal{K}[L, L]$) as well as the Pearson distributions of type I (standard and inverse) and type VI. We will see in the following section that the inverse Pearson distributions of type I find their natural place in this diagram, but offer some theoretical surprises.

4.3 An Original Approach to Characterisation of Three-Parameter Distributions Used for SAR Imagery

Knowing the two elementary two parameter distributions (the gamma distribution and the inverse gamma distribution), it falls natural to make use of these as basic generating functions to obtain a kind of grammar by using elementary operations like the Mellin convolution and the inverse Mellin convolution (One could also have used the Mellin correlation instead of the Mellin convolution while inverting one of the distributions). Assume that we have two distributions p_A and p_B with respective second kind characteristic functions ϕ_A and ϕ_B and log-cumulants $\tilde{\kappa}_A$ and $\tilde{\kappa}_B$. Applying a Mellin convolution or an inverse Mellin convolution will correspond to forming the product or ratio of their second kind characteristic functions, and the sum or difference of the log-cumulants, respectively (See Table 1).

The characteristic functions of the distributions obtained by direct or inverse convolution of the two normalised distributions: the gamma distribution ($\mathcal{G}[1, L]$) or the inverse gamma distribution ($\mathcal{G}[1, M]$), are included in Table 2. From these expressions, and by consulting tables of the Mellin transform (and also using the properties of the transform), it is possible to retrieve the analytical expressions of these distributions without further calculations [14]. Furthermore, while considering only the second and third order second kind cumulants, Table 3 summarises the result obtained by direct or inverse convolution of the two normalised distributions: the gamma distribution ($\mathcal{G}[1, L]$) or the inverse gamma distribution ($\mathcal{G}[1, M]$). Recall that the second and third order second kind cumulants only depend on the shape parameter.

The distributions traditionally used in processing of SAR data are found in this table. These are

- The $\mathcal{K}$ distribution
- The solutions of the Pearson system [6] corresponding to the distributions defined on $\mathbb{R}^+$, that is, in addition to the gamma and inverse gamma distribution, also the type I solutions (also known as the beta distribution) and the type VI solutions (known as the Fisher distribution).

Moreover, uncommon distributions are generated by this algebraic method. It provides:

- A new distribution which is effectively the inverse Pearson distribution of type I, denoted $\mathcal{IPI}[\xi, L, M]$:

$$\mathcal{IPI}[\xi, L, M] = \frac{M}{L\xi} \frac{\Gamma(M)}{\Gamma(L)\Gamma(M-L)} \left(\frac{L\xi}{Mu}\right)^{L+1} \left(1 - \frac{L\xi}{Mu}\right)^{M-L-1} \quad (47)$$

with $u \geq \frac{L\xi}{M}$ and $M \geq L + 1$.

This expression is derived simply by means of the Mellin transform tables [2], since the distribution can be expressed by the relation:

$$\mathcal{IPI}[\xi, L, M] = \mathcal{IG}[\xi, L] \,\hat{\star}^{-1}\, \mathcal{IG}[1, M]$$

whose characteristic function is

$$\xi^{s-1} \frac{\Gamma(L+1-s)}{L^{1-s}\Gamma(L)} \frac{M^{1-s}\Gamma(M)}{\Gamma(M+1-s)}$$

TABLE 2

Second kind characteristic functions of the different distributions obtained by Mellin convolution (direct and inverse) of the gamma distribution $\left(\phi(s) = \mu^{s-1}\frac{\Gamma(L+s-1)}{L^{s-1}\Gamma(L)}\right)$ and inverse gamma distribution $\left(\phi(s) = \mu^{s-1}\frac{\Gamma(L+1-s)}{L^{1-s}\Gamma(L)}\right)$. The distributions whose names are typeset with boldface correspond to new analytical expressions. The second and third order log-cumulants are included in Table 3.

$\hat{\star} \nearrow$	$\mathcal{G}[1, M]$	$\mathcal{IG}[1, M]$
$\mathcal{G}[\mu, L]$	$\mathcal{K}$ distribution $\mu^{s-1}\frac{\Gamma(L+s-1)}{L^{s-1}\Gamma(L)}\frac{\Gamma(M+s-1)}{M^{s-1}\Gamma(M)}$	Pearson VI $\mu^{s-1}\frac{\Gamma(L+s-1)}{L^{s-1}\Gamma(L)}\frac{\Gamma(M+1-s)}{M^{1-s}\Gamma(M)}$
$\mathcal{IG}[\mu, L]$	Pearson VI $\mu^{s-1}\frac{\Gamma(L+1-s)}{L^{1-s}\Gamma(L)}\frac{\Gamma(M+s-1)}{M^{s-1}\Gamma(M)}$	$\mathcal{IK}$ **distribution** $\mu^{s-1}\frac{\Gamma(L+1-s)}{L^{1-s}\Gamma(L)}\frac{\Gamma(M+1-s)}{M^{1-s}\Gamma(M)}$
$\hat{\star}^{-1} \nearrow$	$\mathcal{G}[1, M]$	$\mathcal{IG}[1, M]$
$\mathcal{G}[\mu, L]$	Pearson I $\mu^{s-1}\frac{\Gamma(L+s-1)}{L^{s-1}\Gamma(L)}\frac{M^{s-1}\Gamma(M)}{\Gamma(M+s-1)}$	**Bessel** $\mu^{s-1}\frac{\Gamma(L+s-1)}{L^{s-1}\Gamma(L)}\frac{M^{1-s}\Gamma(M)}{\Gamma(M+1-s)}$
$\mathcal{IG}[\mu, L]$	**Bessel** $\mu^{s-1}\frac{\Gamma(L+1-s)}{L^{1-s}\Gamma(L)}\frac{M^{s-1}\Gamma(M)}{\Gamma(M+s-1)}$	**Inverse Pearson I** $\mu^{s-1}\frac{\Gamma(L+1-s)}{L^{1-s}\Gamma(L)}\frac{M^{1-s}\Gamma(M)}{\Gamma(M+1-s)}$

TABLE 3

Second and third order log-cumulants of the different distributions obtained by Mellin convolution (direct and inverse) of the gamma distribution and the inverse gamma distribution (cf. Table 2). To simplify the presentation, only the second and third order log-cumulants are included in the table. The distributions whose names are typeset in boldface correspond to new analytical expressions.

$\hat{\star} \nearrow$	$\mathcal{G}[1, M]$	$\mathcal{IG}[1, M]$	$\hat{\star}^{-1} \nearrow$	$\mathcal{G}[1, M]$	$\mathcal{IG}[1, M]$
$\mathcal{G}[1, L]$	$\mathcal{K}$ distribution $\Psi(1, L) + \Psi(1, M)$ $\Psi(2, L) + \Psi(2, M)$	Pearson VI $\Psi(1, L) + \Psi(1, M)$ $\Psi(2, L) - \Psi(2, M)$	$\mathcal{G}[1, L]$	Pearson I $\Psi(1, L) - \Psi(1, M)$ $\Psi(2, L) - \Psi(2, M)$	**Bessel** $\Psi(1, L) - \Psi(1, M)$ $\Psi(2, L) + \Psi(2, M)$
$\mathcal{IG}[\mu, L]$	Pearson VI $\Psi(1, L) + \Psi(1, M)$ $-\Psi(2, L) + \Psi(2, M)$	$\mathcal{IK}$ **distribution** $\Psi(1, L) + \Psi(1, M)$ $-\Psi(2, L) - \Psi(2, M)$	$\mathcal{IG}[\mu, L]$	**Bessel** $\Psi(1, L) - \Psi(1, M)$ $-\Psi(2, L) - \Psi(2, M)$	**Inverse Pearson I** $\Psi(1, L) - \Psi(1, M)$ $-\Psi(2, L) + \Psi(2, M)$

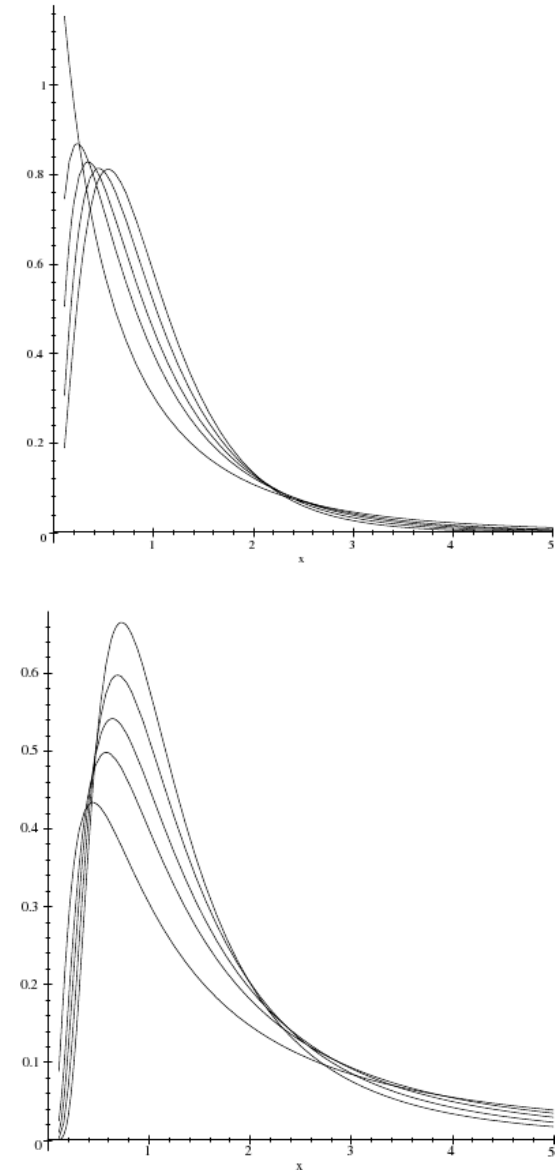


Fig. 7. The $\mathcal{K}$ distribution (top) and the inverse $\mathcal{K}$ distribution (bottom).

where the last expression is found in the Mellin transform tables. Figure 6 allows a comparison between the $\mathcal{IP}$ distribution and the $\mathcal{IPI}$ distribution for the same set of parameters. Recall that the Pearson type I distribution is expressed as:

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

with $u \leq \frac{M}{L\xi}$ and $M \geq L+1$.

Curiously, this distribution is never mentioned in the classical works [9], [19], whereas they characterise the gamma distribution and the inverse gamma distribution separately. Moreover, the inverse Pearson type I distribution has the peculiar property of being localised, in the (β_1, β_2) diagram, between the gamma distribution and the inverse gamma distribution, that is, exactly where the Pearson type VI solution is found. Indeed, the case $M \to \infty$ corresponds to to the inverse gamma distribution and the zone corresponding to the inverse Pearson type I distribution cannot have ambiguities, whereas the (κ_3, κ_2) diagram separate well between the solutions of the Pearson system.

- The inverse $\mathcal{K}$ distribution, which is also uncommon, is expressed as:

$$\mathcal{IK}[\mu, L, M](u) = \frac{1}{L\Gamma(L)M\Gamma(M)} \frac{2}{\mu} \left(\frac{LM\mu}{u}\right)^{\frac{M+L}{2}+1} \times K_{M-L}\left[2\left(\frac{LM\mu}{u}^{\frac{1}{2}}\right)\right] \quad (48)$$

Figure 7 allows a comparison between the $\mathcal{K}$ distribution and the inverse $\mathcal{K}$ distribution for the same set of parameters. As for the $\mathcal{K}$ distribution, the modelling through the Mellin convolution makes it easy to show that the inverse $\mathcal{K}$ distribution tends to an inverse gamma distribution as one of the shape parameters (L or M) goes to to infinity. One thus has a three parameter distribution which is heavy-tailed.

- The combinations $\mathcal{G}\,\hat{\star}^{-1}\,\mathcal{IG}$ and $\mathcal{IG}\,\hat{\star}^{-1}\,\mathcal{G}$ have known analytical solutions that include Bessel functions. However, they are not probability distributions, as the condition $p_x(u) \geq 0$ is not satisfied.

We see that insightful interpretations can be made based on the second and third order log-cumulants, $\tilde{\kappa}_2$ and $\tilde{\kappa}_3$. By the simple analysis of these entities, we can effectively get an idea about the flexibility of a certain distribution compared to the generalised gamma distribution, its inverse, and the other distributions that cover the log-cumulant space: the $\mathcal{K}$ distribution and its inverse, and the solutions of the Pearson system. The analysis of the second and third order log-cumulants can also account for more complex models. We shall see that the same diagram can be used to analyse an additive mixture of gamma distributions, and propose an original and simple practical method to determine the model parameters.

4.4 Additive Mixture of Gamma Distributions

Additive mixtures of gamma distributions are important practical modelling tools (in particular for SAR imagery). Contrarily to the Gaussian case, a unimodal pdf is generally obtained, except when the (two) initial distributions are very different (see Figure 8). However, we will show that there exists a simple solution to determine the parameters of the mixture by analysing this problem aided by second and third order log-cumulants.

Consider the following mixture of gamma distributions:

$$\lambda\mathcal{G}[\mu, L] + \lambda'\mathcal{G}[\mu', L]$$

with $\lambda \geq 0$, $\lambda' \geq 0$ and $\lambda + \lambda' = 1$. In this model, L has the same value for the two gamma distributions.

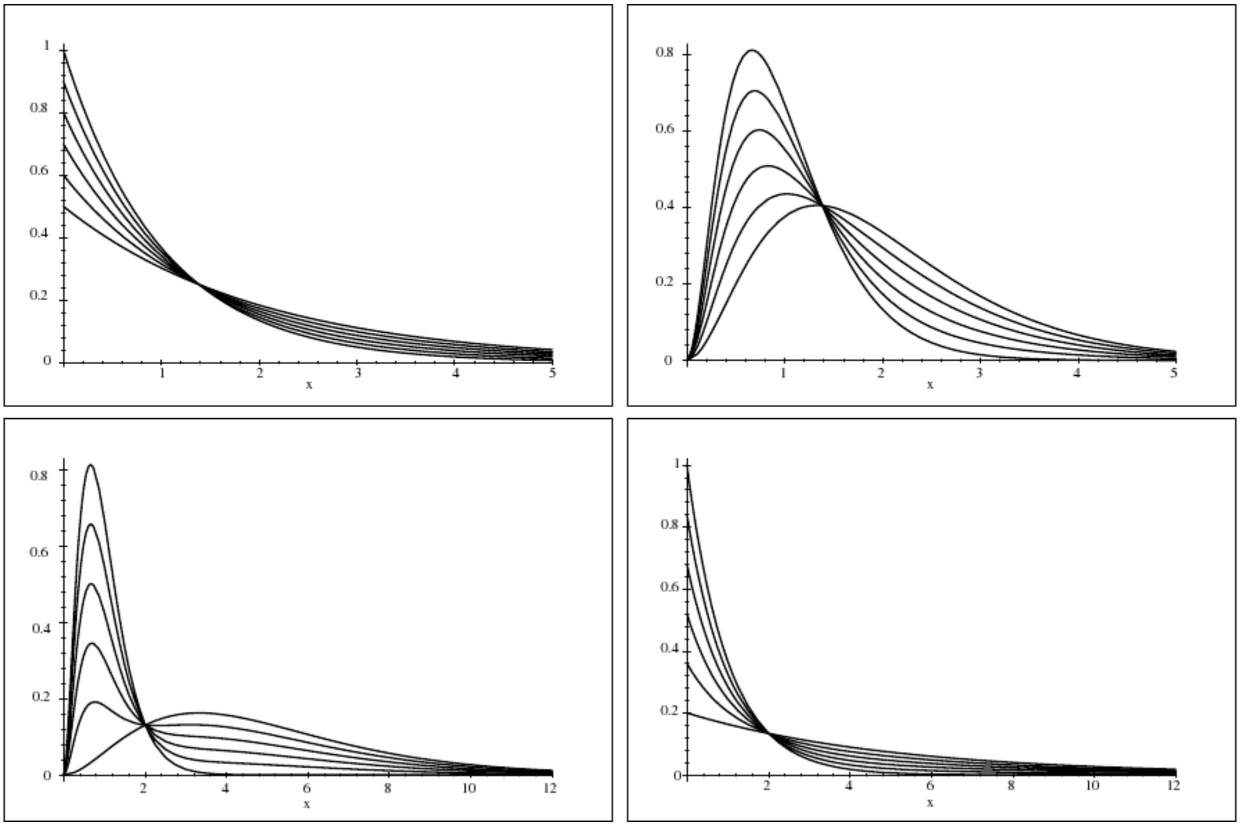


Fig. 8. Examples of additive mixtures of gamma distributions. The left column shows distributions with $L = 1$, and the right column with $L = 3$. In the first row, $\rho = 2$ ($\mu = 1$ and $\mu' = 2$). In the second row, $\rho = 5$ ($\mu = 1$ and $\mu' = 5$). λ takes the values $0, 0.2, 0.4, 0.6, 0.8$ and 1.

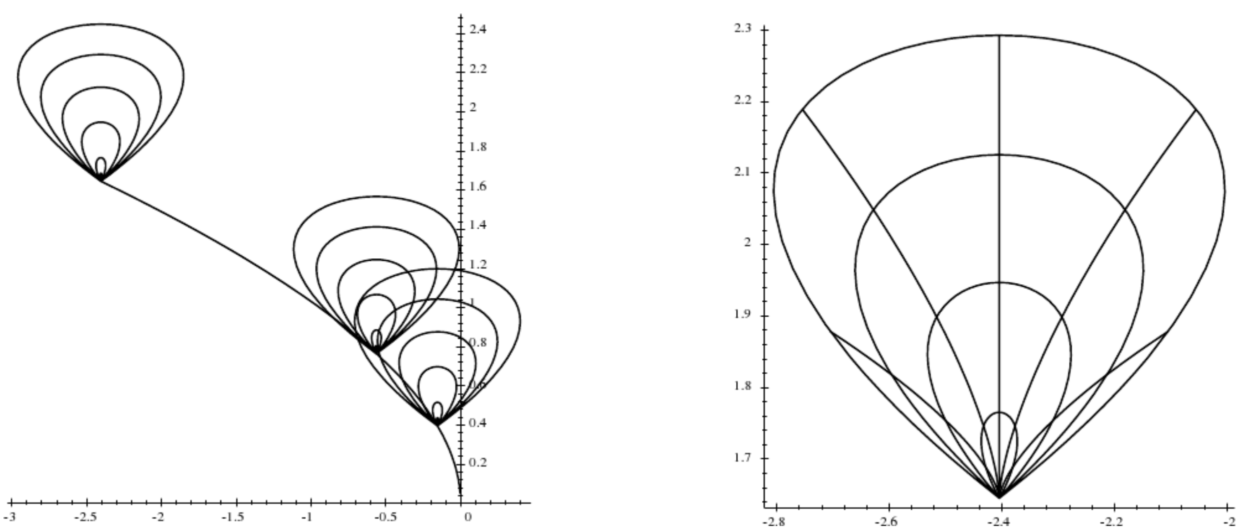


Fig. 9. (κ_2, κ_3) diagram for a mixture of gamma distributions described by the parameters λ (mixing proportion) and ρ (ratio of component means). To the left, for λ varied between 0 and 1, three diagrams are traced out corresponding to several values of ρ for three values of L ($L = 1, 3$ and 5). In the same plot, the gamma distribution is represented by a line which spans $L \in [1, \infty]$. To the right, for a fixed value of $L = 1$, one separately varies ρ between 0 and 5 (five curves, with λ taking the values $0.1, 0.3, 0.5, 0.7$ and 0.9) and λ between 0 and 1 (four closed curves, with ρ taking values $2, 3, 4$ and 5), placing both diagrams in the same figure.

The model can be rewritten by defining the variable ρ such that $\mu' = \rho\mu$, which makes it possible to write the mixture in the following form:

$$\lambda \mathcal{G}[\mu, L] + (1-\lambda)\mathcal{G}[\rho\mu, L] \quad (49)$$

The mixture is then defined by a gamma distribution $\mathcal{G}[\mu, L]$ (corresponding to only one of the mixture components) and two parameters describing the mixture: λ and ρ. The second kind characteristic function is written:

$$\phi(s) = (\lambda + (1-\lambda)\rho^{s-1})\mu^{s-1}\frac{\Gamma(L+s-1)}{L^{s-1}\Gamma(L)}$$

Based on this expression, calculation of the log-cumulants can be carried out directly, giving the following expressions:

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

We remark that starting from the second order, the log-cumulants do not depend on μ, and they have the same values as the standard gamma distribution for the limiting values $\lambda = 0$ and $\lambda = 1$.

We assume that the entity L is known (L can be perceived as a function of the instrument, thus it will be known by the processor). This leads to:

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

Parameters λ and ρ are then given by the solutions of quadratic equation, which gives:

$$\begin{aligned}\lambda &= \frac{1}{2}\left(1 \pm \frac{\overline{\tilde{\kappa}_{x(3)}}}{\sqrt{4\overline{\tilde{\kappa}_{x(2)}}^3 + \overline{\tilde{\kappa}_{x(3)}}^2}}\right) \\ \rho &= e^{\frac{\sqrt{4\overline{\tilde{\kappa}_{x(2)}}^3 + \overline{\tilde{\kappa}_{x(3)}}^2}}{\overline{\tilde{\kappa}_{x(2)}}}}\end{aligned}$$

The evolution of the different parameters in the $(\tilde{\kappa}_3, \tilde{\kappa}_2)$ diagram is shown in Figure 9. It is interesting to notice that the shape of these curves does not depend on L.

If the obtained results are compared with those found in the literature, it is noticed that this approach relies on one assumption only: knowledge of the parameter L, while analyses of mixtures by classical methods require the additional knowledge of μ [20].

4.5 Positive α-Stable Distributions

We will now apply the methodology proposed in this article to the case of a 'heavy-tailed distribution, for which neither the analytical form of the pdf nor moments from a certain order and upwards are known. This prohibits the method of moments.

A positive α-stable distribution has a pdf characterised by two parameters: α and γ. It cannot in general be defined, other than by its characteristic function $\Phi(\nu)$, which is written (according to Pierce [17]) as:

$$\Phi(\nu) = e^{-\gamma|\nu|^\alpha\left(1 + j\,\mathrm{sgn}(\nu)\tan\left(\frac{\alpha\pi}{2}\right)\right)}$$

with

$$\mathrm{sgn}(\nu) = \begin{cases} 1, & \nu > 0 \\ 0, & \nu = 0, \\ -1, & \nu < 0 \end{cases} \qquad 0 < \alpha < 1, \quad \gamma > 0\,.$$

Except for certain particular values of α, the analytical expression of the distribution is not known.

One can nevertheless express the moments of this distribution (including fractional ones) as:

$$m_\nu = \frac{\gamma^{\frac{\nu}{\alpha}}\sin(\pi\nu)\Gamma(\nu+1)\left(1 + \left(\tan\left(\frac{\pi\alpha}{2}\right)\right)^2\right)^{\frac{\nu}{2\alpha}}}{\alpha\sin\left(\frac{\pi\nu}{\alpha}\right)\Gamma\left(1 + \frac{\nu}{\alpha}\right)} \quad (50)$$

These moments are only defined for $\nu < \alpha < 1$, which means that even the first moment is not defined. This is evidently a heavy-tailed distribution.

It is nevertheless possible, by an analytical continuation, to derive the second order characteristic function, which is written:

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

It is seen that this function is well defined in a complex neighbourhood around the value $s = 1$. The existence theorem in Section 2.5 thus confirms that the log-moments and log-cumulants of all orders exist, whereas the moments of orders $\nu \geq \alpha$ are not defined.

Even though the analytical form is rather complicated, it is still possible to obtain simple expressions of the log-cumulants. Note that these expressions are only analytical continuations because the gamma functions in the derivatives of $\phi(s)$ have discontinuities at $s = 1$, a value at which they must be evaluated when calculating the log-cumulants. It is then necessary to study the limit at $s = 1$ in order to obtain the analytical expressions. The following results were established with assistance of mathematical computation software Maple, as a result of lengthy analytical developments:

$$\begin{aligned}\tilde{\kappa}_1 &= \frac{(1-\alpha)\Psi(1)}{\alpha} + \frac{-\log\left(\cos\left(\frac{\pi\alpha}{2}\right)\right)}{\alpha} + \frac{\log\gamma}{\alpha} \\ \tilde{\kappa}_2 &= \frac{(1-\alpha^2)}{\alpha^2}\Psi(1,1) \\ \tilde{\kappa}_3 &= \frac{\alpha^3 - 1}{\alpha^3}\Psi(2,1)\end{aligned} \quad (51)$$

These expressions, that eventually appear as rather simple, illustrate well the strength of our new approach. The two parameters of the distribution are easily derived since:

- The parameter α can be estimated from the second order log-cumulant as:

$$\alpha = \sqrt{\frac{\Psi(1,1)}{\Psi(1,1) + \tilde{\kappa}_2}}$$

- When α is known, γ can be obtained from the first log-cumulant as:

$$\gamma = e^{\alpha \tilde{\kappa}_1 - (1-\alpha)\Psi(1) + \log\left(\cos\left(\frac{\pi\alpha}{2}\right)\right)}$$

- By combining the previous expressions, one can also write γ in terms of $\tilde{\kappa}_1$ and $\tilde{\kappa}_1$:

$$\gamma = \exp\left(\sqrt{\frac{\Psi(1,1)}{\Psi(1,1)-\tilde{\kappa}_2}}\tilde{\kappa}_1 - \left[1 - \sqrt{\frac{\Psi(1,1)}{\Psi(1,1)-\tilde{\kappa}_2}}\right]\Psi(1) + \log\left[\cos\left(\frac{\pi}{2}\sqrt{\frac{\Psi(1,1)}{\Psi(1,1)-\tilde{\kappa}_2}}\right)\right]\right)$$

We note that the Mellin transform sheds, on the theoretical side, a new and recent light on the heavy-tailed distributions [21].

4.6 A Generalisation of the Rayleigh Distribution

Another example drawn from the α-stable distributions inspired Kuruoğlu and Zerubia to propose a generalisation of the Rayleigh distribution [18]. The pdf has two parameters (α and γ) and its analytical expression is given by the following integral equation:

$$p(u) = u \int_0^{\infty} v e^{-\gamma v^{\alpha}} J_0(uv)\, dv \qquad (52)$$

where J_0 is the Bessel function of the second kind. This distribution falls into the heavy-tailed category, since its moments are not defined from a certain order and upwards, with this order given as: $\min(\alpha, 2)$.

To calculate its second kind characteristic function, two approaches are possible:

- The first search for the Mellin transform of this expression [14] led to the following result:

$$\phi(s) = \frac{2^s \Gamma\left(\frac{s+1}{2}\right) \gamma^{\frac{s-1}{\alpha}} \Gamma\left(\frac{1-s}{\alpha}\right)}{\Gamma\left(\frac{1-s}{2}\right)\alpha} \qquad (53)$$

- A second approach consists of rewriting (52) on the form of a Mellin correlation:

$$p(u) = u\left(J_0(u)\,\hat{\otimes}\left(e^{-\gamma u^{\alpha}}\right)\right)$$

  By using the property in (22), (53) is immediately retrieved.

It can be noted that at $s = 1$, the second kind characteristic function of this distribution goes in the limit to the value 1, since

$$\lim_{s\to 1} \frac{\Gamma\left(\frac{1-s}{\alpha}\right)}{\Gamma\left(\frac{1-s}{2}\right)} = \frac{\alpha}{2}$$

It is thus a valid pdf.

In the vicinity of $s = 1$, this function is defined for $s < 1 + \min(\alpha, 2)$ and for $s > -1$. It is thus well defined in a vicinity of $s = 1$, hence it is legitimate to calculate its log-moments and log-cumulants. Note that the case

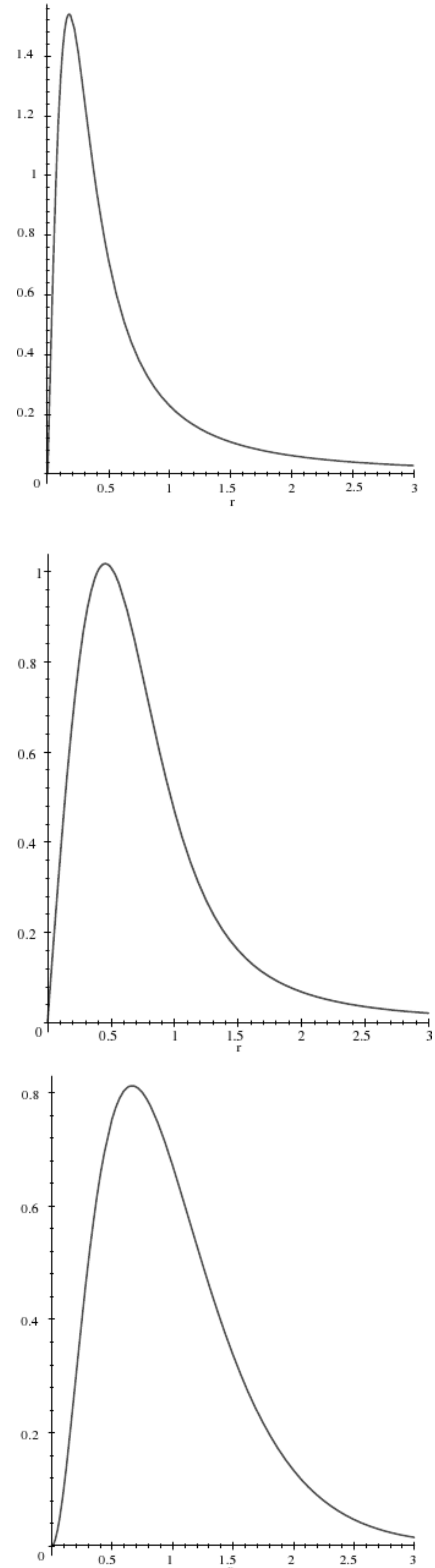


Fig. 10. Generalised Rayleigh distribution [18]: $\alpha = 1$, $\alpha = 1.5$ and $\alpha = 2$.

$\alpha = 2$ gives the Rayleigh distribution and the case $\alpha = 1$ gives the distribution

$$p(u) = \frac{\gamma u}{(u^2+\gamma^2)^{\frac{3}{2}}} .$$

Figure 10 shows this distribution for $\alpha = 1$, $\alpha = 1.5$ and $\alpha = 2$ (the Rayleigh distribution).

Although the analytical form is rather complicated, it is possible to formulate the first and second order log-cumulants of the probability distribution. Again, the expressions obtained are analytical continuations because of discontinuities in the gamma functions. Also in this case, Maple was used to establish following expressions:

$$\tilde{\kappa}_1 = -\Psi(1)\frac{1-\alpha}{\alpha} + \log\left(2\gamma^{\frac{1}{\alpha}}\right)$$
$$\tilde{\kappa}_2 = \frac{\Psi(1,1)}{\alpha^2} .$$

The equation system obtained is easy to handle. The distribution parameters are easily retrieved from the first two log-cumulants:

$$\alpha = \sqrt{\frac{\Psi(1,1)}{\tilde{\kappa}_2}}$$
$$\gamma = e^{\alpha\tilde{\kappa}_1 - \Psi(1)(1-\alpha)} .$$

It is verified that for $\alpha = 2$, the log-cumulants of the Rayleigh distribution are retrieved (with $\mu = 2\sqrt{\gamma}$):

$$\tilde{\kappa}_1 = \frac{1}{2}\Psi(1) + \log(2\sqrt{\gamma})$$
$$\tilde{\kappa}_2 = \frac{1}{4}\Psi(1,1) .$$

5 Parameter estimation

The proposal of a new methodology to evaluate the parameters of a probability distribution requires us to compare the results with those obtained by traditional methods in a realistic setting where N samples are available. In order to decide which method is the preferred one, it is important to establish the variance of the estimators. The goal of this section is to carry out an exhaustive comparison for a schoolbook example: the gamma distribution, for which the method of moments, the method of log-moments, and the method of maximum likelihood are applicable.

5.1 Traditional Methods: Method of Moments Estimation and Maximum Likelihood Estimation

5.1.1 Experimental Framework

Assume that a probability distribution is described by p parameters: α_j, $j \in [1,p]$. Estimation of the parameters describing this distribution is commonly performed by the two following approaches:

- The *method of moments* (MoM) consists of calculating the sample moments of order 1 to p in order to obtain a system of p equations in p unknown pdf parameters. If N samples are available, x_i, $i \in [1,N]$, the rth order sample moment is expressed simply as

$$m_r = \frac{1}{N}\sum_{i=1}^{N} x_i^r .$$

  In order to determine p parameters, it is necessary that all moments up to order p exist, which can pose a problem for instance for distributions with heavy tails. It is also possible to use fractional moments (like FLOM [3]), or lower (and even negative) order moments, whose possible existence is justified in section 2.2 [4], to obtain an equation system which can be solved. Note, however, that the expressions sometimes prove impossible to invert analytically, and the system may also be difficult to invert numerically.
- The *maximum likelihood* approach consists of regarding the N samples x_i as N independent realisations of the distribution which one seeks to estimate, so that they maximise the expression

$$\prod_{i=1}^{N} p_x(x_i)$$

  or, equivalently,

$$\sum_{i=1}^{N} \log(p_x(x_i)) .$$

  With these expressions representing a maximum, calculation of partial derivatives for each parameter then makes it possible to obtain another system of p equations in p unknowns:

$$\frac{\partial\left(\sum_{i=1}^{N}\log(p_x(x_i))\right)}{\partial\alpha_j} = 0 . \qquad (54)$$

  The solution relies on the existence of the partial derivatives, which can pose a problem, as for the $\mathcal{K}$ distribution [16].

5.1.2 Estimator Variance

With several applicable methods available, we must compare them to select the one which is likely to give the user the most reliable results. A natural approach is to seek the method which provides minimum variance for the estimator of a given parameter, knowing that one has a finite number of N samples.

It is known that for the distributions of the exponential family, maximum likelihood estimators attain the Cramer-Rao boundary. Provided that the p partial derivatives in Eq. (54) can be calculated analytically, and that the system of equations can be solved, one obtains p estimators whose variances are minimal. However, many existing distributions (such as the $\mathcal{K}$ distribution) do not have analytical expressions for all partial derivatives, which then renders the method of maximum likelihood inadequate.

In this case, the use of the method of moments is justified, even if the estimator variance thus obtained is higher. The variance of estimators obtained by the method of moments can be derived through an approach suggested by Kendall [19]. Let m_1 and m_2 be the estimates of the first two moments, and $g(m_1, m_2)$ a function depending only on these two entities. We seek to calculate the variance of the function $g(m_1, m_2)$ by linearising it and writing it as a first-order expansion around the values of the theoretical moments, $m_{0,1}$ and $m_{0,2}$:

$$g(m_1,m_2) = g(m_{0,1},m_{0,2}) + (m_1 - m_{0,1})\frac{\partial g}{\partial m_1}(m_{0,1},m_{0,2}) + (m_2 - m_{0,2})\frac{\partial g}{\partial m_2}(m_{0,1},m_{0,2})\,.$$

After having verified that the $\partial g/\partial m_i$ are not both zero in the point $(m_{0,1}, m_{0,2})$, the variance of g is established as the quadratic error between $g(m_1, m_2)$ and $g(m_{0,1}, m_{0,2})$ due to the following formula [19, Eq. (10.12)]:

$$\begin{aligned}
&\mathrm{Var}\{g(m_1,m_2)\} \\
&= \mathrm{E}\left\{[g(m_1,m_2) - g(m_{0,1},m_{0,2})]^2\right\} \\
&= \mathrm{E}\left\{\left[(m_1 - m_{0,1})\frac{\partial g}{\partial m_1}(m_{0,1},m_{0,2}) + (m_2 - m_{0,2})\frac{\partial g}{\partial m_2}(m_{0,1},m_{0,2})\right]^2\right\} \\
&= \frac{\partial g}{m_1}(m_{0,1},m_{0,2})^2\,\mathrm{Var}\{m_1\} + \frac{\partial g}{m_2}(m_{0,1},m_{0,2})^2\,\mathrm{Var}\{m_2\} \\
&\quad + 2\frac{\partial g}{\partial m_1}(m_{0,1},m_{0,2})\frac{\partial g}{\partial m_2}(m_{0,1},m_{0,2}) \times \mathrm{Cov}\{m_1,m_2\}\,.
\end{aligned} \tag{55}$$

The method can obviously be generalised to functions utilising moments m_i of order i. The definition of the covariance matrix allows us to write:

$$\mathrm{Var}\{m_i\} = \frac{1}{N}(m_{2i} - m_i^2)$$
$$\mathrm{Cov}\{m_i,m_j\} = \frac{1}{N}(m_{i+j} - m_i m_j)$$

5.2 Method of Log-Moments

We propose in this article a new method for analysis of pdfs defined on $\mathbb{R}^+$ based on log-moments and log-cumulants. We will see in this section how to implement it and how to calculate the variance of the estimators obtained.

5.2.1 Description

The method of log-moments (MoLM) consists of calculating estimates of log-moments and log-cumulants in order to obtain a system of p equations in p unknowns (the parameters of the pdf). Assume that we have N samples $x_i, i \in [1, N]$ from the distribution to be estimated. The estimate of the pth order log-moment is expressed simply as

$$\tilde{m}_p = \frac{1}{N}\sum_{i=1}^{N}\log x_i^p\,.$$

To determine p parameters, it is necessary to check in advance that the log-moments up till order p exist. This is in general true, as stated by the theorem of existence, which has been verified for the distributions generally used in signal and image processing.

5.2.2 Estimator Variance

Since we use a logarithmic scale, the criterion of the quadratic error (applied in Eq. (55)), is replaced by another criterion which we will call "normalised quadratic error", E_{nq}, which is in fact the quadratic error calculated on a logarithmic scale:

$$\mathrm{E}_{nq} = \mathrm{E}\left\{\left(\log\left(\frac{x}{y}\right)\right)^2\right\}\,.$$

In the same spirit, we introduce the second kind variance and covariance, $\tilde{\mathrm{Var}}$ and $\tilde{\mathrm{Cov}}$, on the form

$$\tilde{\mathrm{Var}}\{\tilde{m}_i\} = \mathrm{E}\left\{\left[(\log x)^i - \tilde{m}_i\right]^2\right\} = \frac{1}{N}\left(\tilde{m}_{2i} - \tilde{m}_i^2\right)$$
$$\tilde{\mathrm{Cov}}\{\tilde{m}_i,\tilde{m}_j\} = \mathrm{E}\left\{\left[(\log x)^i - \tilde{m}_i\right]\left[(\log x)^j - \tilde{m}_j\right]\right\} = \frac{1}{N}(\tilde{m}_{i+j} - \tilde{m}_i\tilde{m}_j)$$

where $\tilde{m}_i$ is the ith order log-moment.

With this new approach, and taking the preceding step as starting point, let the function g be expressed in terms of the first two estimated log-moments as $g(\tilde{m}_i, \tilde{m}_2)$. Then g can be expanded around the first two theoretical log-moments, $\tilde{m}_{0,1}$ and $\tilde{m}_{0,2}$, as

$$g(\tilde{m}_1,\tilde{m}_2) = g(\tilde{m}_{0,1},\tilde{m}_{0,2}) + (\tilde{m}_1 - \tilde{m}_{0,1})\frac{\partial g}{\partial \tilde{m}_1}(\tilde{m}_{0,1},\tilde{m}_{0,2}) + (\tilde{m}_2 - \tilde{m}_{0,2})\frac{\partial g}{\partial \tilde{m}_2}(\tilde{m}_{0,1},\tilde{m}_{0,2})\,.$$

After verifying that the $\partial g/\partial \tilde{m}_i$ are not both zero in $(\tilde{m}_{0,1}, \tilde{m}_{0,2})$, the variance of g is established by the same formula applied in the previous section.

$$\begin{aligned}
&\mathrm{Var}\{g(\tilde{m}_1,\tilde{m}_2)\} = \mathrm{E}\left\{[g(\tilde{m}_1,\tilde{m}_2) - g(\tilde{m}_{0,1},\tilde{m}_{0,2})]^2\right\} \\
&= \mathrm{E}\left\{\left[(\tilde{m}_1 - \tilde{m}_{0,1})\frac{\partial g}{\partial \tilde{m}_1}(\tilde{m}_{0,1},\tilde{m}_{0,2}) + (\tilde{m}_2 - \tilde{m}_{0,2})\frac{\partial g}{\partial \tilde{m}_2}(\tilde{m}_{0,1},\tilde{m}_{0,2})\right]^2\right\} \\
&= \frac{\partial g}{\tilde{m}_1}(\tilde{m}_{0,1},\tilde{m}_{0,2})^2\tilde{\mathrm{Var}}\{\tilde{m}_1\} + \frac{\partial g}{\tilde{m}_2}(\tilde{m}_{0,1},\tilde{m}_{0,2})^2\tilde{\mathrm{Var}}\{\tilde{m}_2\} \\
&\quad + 2\frac{\partial g}{\partial \tilde{m}_1}(\tilde{m}_{0,1},\tilde{m}_{0,2})\frac{\partial g}{\partial \tilde{m}_2}(\tilde{m}_{0,1},\tilde{m}_{0,2}) \times \tilde{\mathrm{Cov}}\{\tilde{m}_1,\tilde{m}_2\}\,.
\end{aligned}$$

As in the previous, this method can obviously be generalised to functions of the moments $\tilde{m}_i$ of unspecified order i.

5.3 The Gamma Distribution Case

We will use the gamma distribution as an example to compare the available methods. This distribution is not heavy tailed, thus the method of moments can be used, as well as the method of log-moments. The partial derivatives with respect to the parameters are known, which makes it possible to apply maximum likelihood estimation.

5.3.1 Variance of the Gamma Distribution Parameter Estimators with the Method of Moments

The method of moments (MoM) utilises the first two moments to deduce estimates of L and μ (Eqs. (30) and (31)). The method of Kendall, presented in Section 5.1.2, gives the following variance for the estimators of μ and L:

$$\mathrm{Var}_{MoM}\{\hat{\mu}\} = \frac{\mu^2}{NL} \tag{56}$$

$$\mathrm{Var}_{MoM}\{\hat{L}\} = \frac{2L(L+1)}{N} \tag{57}$$

5.3.2 Variance of the Gamma Distribution Parameter Estimators with the Method of Log-Moments

The parameter L is derived from (33) as

$$\Psi(1, L) = \tilde{\kappa}_{x(2)}$$

which can be rewritten as a function of $(\tilde{m}_1, \tilde{m}_2)$:

$$\Psi(1, L) = \tilde{m}_2 - \tilde{m}_1^2 .$$

One then carries out the limited expansion proposed in the previous, which requires the use of implicit differentiation. Although the expression brings into play the first to fourth order log-moments, the result can be simplified and we obtain:

$$\mathrm{Var}_{MoLM}\{\hat{L}\} = \frac{1}{N}\frac{\Psi(3, L) + 2\Psi(1, L)^2}{\Psi(2, L)^2} . \tag{58}$$

Figure 11 (left panel) presents the ratio of the standard deviation for the variance of MoLM estimate of $\hat{L}$ to the standard deviation for the variance of the MoM estimate of $\hat{L}$. The whole motivation for using the new method is evident for low values of L, where the improvement approaches 30%. When the variance of the different estimators is fixed, this results in the same amount of shrinking of the analysis window, and therefore a better spatial localisation of the estimate.

For the parameter μ, the calculation is much more elaborate, and we finally arrive at the following expression, whose interpretation is not simple, but which can easily be implemented numerically:

$$\begin{aligned}\mathrm{Var}_{MoLM}\{\hat{L}\} &= -\frac{1}{N}\frac{\mu^2}{L^2\Psi(2, L)^2} \\ &\times \big[2\Psi(1, L)L\Psi(3, L) - \Psi(1, L)^2L^2\Psi(3, L) \\ &\quad + 4\Psi(1, L)^3L - 2\Psi(1, L)^4L^2 - 2L\Psi(2, L)^2 \\ &\quad + \Psi(1, L)L^2\Psi(2, L)^2 - 2\Psi(1, L)^2 - \Psi(3, L)\big]\end{aligned}$$

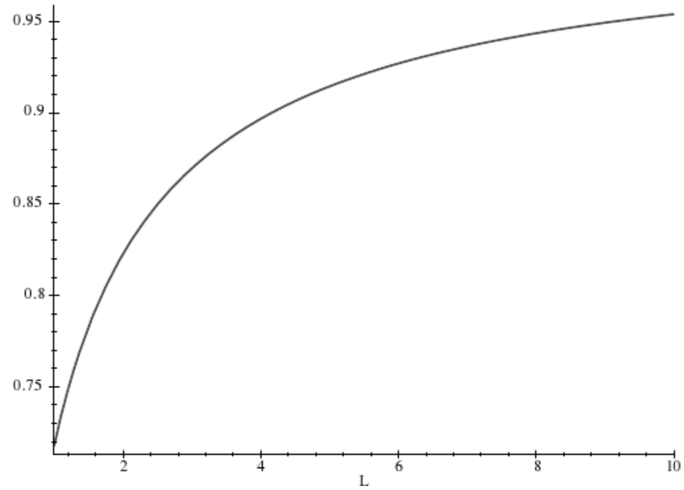


Fig. 12. Gamma distribution: Comparison of the variance of the estimator for μ by the method of log-moments with the method of moments. The curve represents the ratio of the standard deviations for values of L between 1 and 10.

Also Figure 12 presents the ratio of the standard deviation of $\hat{\mu}$ calculated by the MoLM to the standard deviation of $\hat{\mu}$ calculated by the MoM.

It can be noted that the MoM provides better results for low values of L. Recall moreover that this is also the maximum likelihood estimator and thus attains minimum variance (i.e. the Cramer-Rao bound).

5.3.3 Variance of the Gamma Distribution Parameter Estimators with the Method of Lower Order Moments

The existence of the second kind characteristic function for values of s lower than 1 justifies the use of the method of lower order moments (MoLOM), i.e. negative ones. In the case of the gamma distribution, it is known that the lower order moments exist for $\nu > -L$. For a given value of ν it is verified that $\nu > -L$ and using the three moments μ_ν, $\mu_{\nu+1}$ and $\mu_{\nu+2}$ it is easy to show that $\hat{\mu}$ and $\hat{L}$ can be derived from the relation:

$$\hat{\mu} = \frac{\hat{m}_{\nu+1}}{\hat{m}_\nu}(1+\nu) - \nu\frac{\hat{m}_{\nu+2}}{\hat{m}_{\nu+1}}$$

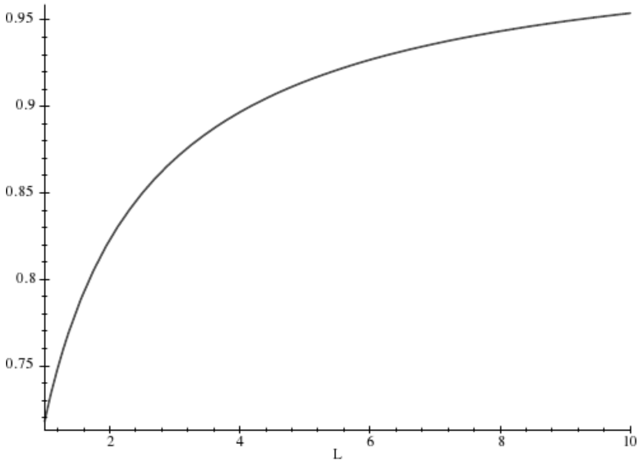
$$\hat{L} = \frac{1}{\hat{m}_\nu\left(\frac{\hat{m}_{\nu+2}}{\hat{m}_{\nu+1}}\right) - 1} - \nu$$

For $\nu = 0$, this reduces to the MoM (Eqs. (30) and (31)).

The variances of the estimators for μ and L can be established by the method of Kendall, used in Section 5.3.1 (for the MoM). For L, the following expression is obtained:

$$\begin{aligned}\mathrm{Var}_{MoLOM}\{\hat{L}\} &= \frac{1}{N}\frac{\Gamma(L)\Gamma(2\nu + L)}{\Gamma(\nu + L)^2} \\ &\times \left(2L(L+1) + \nu[4L(\nu+2) + (\nu+4)(\nu+1)^2]\right)\end{aligned} \tag{59}$$

The problem with this relation is that it has a minimum for ν, which cannot be expressed explicitly as a function of L. The optimal values of ν must be calculated numerically. Table 4 gives some values of ν as a function



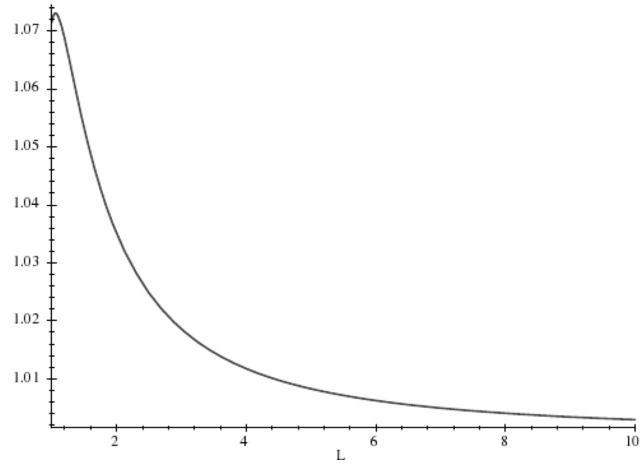


Fig. 11. To the left, a comparison of the variance of the estimator for L by the method of log-moments and the method of moments. To the right, a comparison of the variance of the estimators of L by the method of log-moments and the method of lower order moments (only for the value of $\nu = -0.35$). The curves represent the standard deviation for values of L between 1 and 10.

TABLE 4
Gamma distribution estimated with the method of lower order moments. Optimal values of the parameter ν that minimises the variance of $\hat{L}$ as a function of L

L	ν_{opt}
1	-0.35
2	-0.44
3	-0.56
4	-0.59

of L. When information about L is absent, the choice of $\nu = -0.35$ seems to be a good compromise.

Figure 11 (right panel) presents the ratio of the standard deviation of $\hat{L}$ calculated by the MoLM to the standard deviation of $\hat{L}$ calculated by the MoLOM with $\nu = -0.35$. It is interesting to note that the MoLOM is slightly better than the MoLM. Nevertheless, if one wants to fully utilise this method, then one must know L to be able to choose the optimal value of ν. As the difference is altogether rather small, we promote the MoLM because it does not require us to determine a parameter in order to make optimal use of the method.

However, it is easily shown that minimum variance is obtained for $\nu = 0$, which is less than astonishing since this value corresponds to the maximum likelihood estimator.

5.3.4 Summary

We propose to summarise these results by posting in Table 5 the optimal window dimension for these three methods when we seek to reach an error of 10% for the estimate of the shape parameter L (i.e. the standard deviation is 10% of the value to be estimated). For the parameter μ, only the method of moments (which coincides with the maximum likelihood estimator) is used.

TABLE 5
Number of samples (and examples of the analysis window) needed to estimate the parameters L and μ of a gamma distribution with 10% error. The methods used are, for the shape parameter L, the method of moments (MoM), the method of lower order moments (MoLOM) with $\nu = -0.35$, and the method of log-moments (MoLM). The Cramer-Rao bound (CRB) is calculated by the means of the Fisher information matrix. For μ, only the MoM is used.

Gamma distribution					
L	$\hat{L}$				$\hat{\mu}$
	MoM	MoLOM	MoLM	CRB	MoM
1	400 20×20	179 13×13	206 14×14	155 12×12	100 10×10
2	300 17×17	189 14×14	203 14×14	172 13×13	50 7×7
3	267 16×16	194 14×14	202 14×14	180 13×13	33 6×6
5	240 15×15	197 14×14	201 14×14	187 14×14	20 4×4
10	220 14×14	199 14×14	200 14×14	194 14×14	10 3×3

We first remark that for an identical relative error (10%), the estimate of L requires much more samples

TABLE 6
Number of samples (and examples of the analysis windows) needed to estimate the parameters σ and μ of a Gaussian distribution with 10% error.

Gaussian distribution		
σ	$\hat{\sigma}$	$\hat{\mu}$
1	200 14×14	100 10×10
0.707	200 14×14	50 7×7
0.577	200 14×14	33 6×6
0.447	200 14×14	20 4×4
0.316	200 14×14	10 3×3

than the estimate of μ. Secondly, the estimate of μ requires much less samples when L is large, i.e. the distribution is localised. Lastly, it is interesting to note a characteristic feature of the MoLM: It requires about the same number of samples for all values of L, whereas the MoM requires a much higher number of samples when L is small. From this, two remarks can be made:

- It can be shown that the variance of $\hat{L}$ for the MoLM (Eq. (58)) is almost quadratic in L:

$$\frac{\Psi(3, L) + 2\Psi(1, L)^2}{\Psi(2, L)^2} \simeq 2L^2$$

  Thus, if a constant relative error is sought, the number of samples is independent of L.
- It is interesting to analyse the same problem for the Gaussian distribution $\mathcal{N}[\mu, \sigma^2]$. It is easy to show that the variances of the estimators of μ and σ do not depend on σ for the MoM. They are written:

$$\text{Var}_{\mathcal{N}}(\mu) = \frac{\sigma^2}{N}$$
$$\text{Var}_{\mathcal{N}}(\sigma) = \frac{\sigma^2}{2N}$$

  By choosing Gaussian distributions with $\mu = 1$, the values of σ become comparable and an identical criterion, and the required window sizes can be calculated. These are included in Table 6.

It is seen that the MoM needs a constant number of samples to estimate the shape parameter σ, an analogy to the property of the MoLM for the gamma distribution.

To achieve this analysis, we calculate Fisher's information matrix for the gamma distribution[4]:

$$\begin{bmatrix} \frac{L}{\mu^2} & 0 \\ 0 & \Psi(1, L) - \frac{1}{L} \end{bmatrix},$$

which allows the calculation of the Cramer-Rao bound, given in Table 5.

5.4 The Mixture of Gamma Distribution Case

The analytical calculation of the variance of the estimators in the mixture of gamma distributions case described in Section 4.4 does not pose any problem, except for the apparent complexity of the expressions obtained, whose length prohibits us from including them in a publication. Another possibility would be to assess them by numerical evaluation.

Table 7 presents the standard deviations of λ and ρ if the analysis is carried out in neighbourhood of 100 samples (a 10×10 window), for various values of ρ and λ.

Table 8 presents the optimal dimension of a square window which guarantees a maximum of 10% estimation error (where the error is defined as the ratio of the standard deviation to the estimated value). Note that for $L = 1$, a large window size is required, which is not surprising when recalling Figure 8 showing that a mixture of gamma distributions is generally unimodal.

6 Conclusions

Second kind statistics seem to be an innovative and powerful tool for the study of distributions defined on $\mathbb{R}^+$. The analytical formulation of the log-moments and the log-cumulants is indeed particularly simple and easy to exploit. At least, this is true for the examples presented in this article, whereof some, such as the mixture distributions, are not commonplace. Moreover, the variance of the estimators thus defined approaches the minimal values reached by the maximum likelihood method, while avoiding some of the analytical pitfalls. This approach shows great potential in certain applications in SAR image processing (such as the characterisation of an optimal homomorphic filter [11]). One can reasonably question why this approach, in all its simplicity, has not been proposed before. Several reasons can be called upon:

- The first is based on the observation that a Mellin transform of a pdf is only a Fourier transform of the same pdf taken on a logarithmic scale. Even if this step is perfectly justified on the theoretical level, all the possible advantages of moving into the Mellin domain remain hidden, such as the use of existing tables of known Mellin transforms, or the direct use of the log-moments and log-cumulants that produce better estimates of the distribution parameters.

4. The diagonal form of this matrix justifies *a posteriori* the analytical expression of the gamma distribution that we chose, which differs slightly from the one found in reference book like [9], [19].

TABLE 7

Standard deviation (SD) for the estimates of λ and ρ in the case of a mixture of gamma distributions. The size of the analysis window is 10×10. Since the SD is inversely proportional to the square root of the number of samples in the analysis window, the table can serve to determine the optimal window after a maximum error has been set.

$\rho = 2$			$\rho = 5$			$\rho = 10$		
λ	SD_λ	SD_ρ	λ	SD_λ	SD_ρ	λ	SD_λ	SD_ρ
.1	.553	3.874	.1	.071	2.638	.1	.043	3.867
.2	.617	1.967	.2	.083	1.606	.2	.052	2.481
.3	.676	1.235	.3	.092	1.183	.3	.059	1.918
.4	.730	.784	.4	.099	.925	.4	.063	1.591
.5	.780	.455	.5	.105	.749	.5	.065	1.380

TABLE 8

Optimal window size for a relative error of 10% in λ and ρ.

$\rho = 2$			$\rho = 5$			$\rho = 10$		
λ	SD_λ	SD_ρ	λ	SD_λ	SD_ρ	λ	SD_λ	SD_ρ
.1	553×553	194×194	.1	71×71	53×53	.1	43×43	39×39
.2	309×309	98×98	.2	41×41	32×32	.2	26×26	25×25
.3	225×225	62×62	.3	31×31	24×24	.3	20×20	19×19
.4	182×182	39×39	.4	25×25	19×19	.4	16×16	16×16
.5	156×156	23×23	.5	21×21	15×15	.5	13×13	14×14

- The analysis of the product model, which is reserved for particular processes like coherent imaging, has not received the same strong attention as the additive signal model. The philosophy adopted for the study of the product model has too often consisted of transformation into logarithmic scale, in order to use the known tools for the additive model. This reductional step quickly pose problems, as it requires large control of the analytical expressions thus obtained. It is probably the reason why non-experts have written off other distributions than the gamma distribution and the inverse gamma distribution, such as the $\mathcal{K}$ distribution, for instance.
- Finally, the Mellin transform has been completely ignored. Its applications has been confined to certain specialised applications, which has unfortunately prevented diffusion of the method beyond the field of study (e.g., radar and sonar signals, number theory, ultrasound propagation in heterogeneous media, the Fourier-Mellin transform in image processing). Even if certain pieces of work, old [15] as well as recent [21], [22] ones, have shown its applicability in the field of probability, its use has been very restricted. Therefore, few people know the fundamental properties, or even the exact definition.

The unfortunate consequence of the confidentiality is that few research groups have worked on the subject. Therefore, powerful and sufficiently general numerical implementations of the analytical transform are still missing. These would make it possible to consider numerical deconvolutions of the probability distributions described by a Mellin convolution, and thus to recover significant parameters of a SAR scene [23].

REFERENCES

[1] M. Métivier, *Notions Fundamentales de la Théorie des Probabilités*. Dunod, 1972.

[2] S. Colombo, *Les Transformations de Mellin et de Hankel*. Centre National de la Recherche Scientifique, 1959.

[3] M. Shao and C.-L. Nikias, "Signal processing with fractional lower order moments," *Proc. of the IEEE*, vol. 81, no. 7, pp. 986–1010, Jul. 1993.

[4] J.-M. Nicolas, A. Maruani, and R. Badeau, "Les moments d'ordre inferieur: Principles et application au filtrage des images RSO," in *RFIA'00*, Paris, France, Jan. 2000.

[5] J.-M. Nicolas, M. Sigelle, C. Thuillier, and F. Tupin, "Images de radar à ouverture synthétique: Transformée de Mellin et multirésolution," in *GRETSI'97*, Grenoble, France, Sep. 1997.

[6] J.-M. Nicolas and A. Maruani, "Speckle well modeled by Mellin transform," in *PIERS'98*, Nantes, France, Jul. 1998.

[7] H. Cartan, *Théorie Élémentaire des Fonctions Analytiques d'Une ou Plusieurs Variables Complexes*. Hermann, 1961.

[8] E. Lukacs, *Developments in Characteristic Functions Theory*. Charles Griffin, 1983.

[9] N. Johnson and S. Kotz, *Continuous Univariate Distributions 1*. John Wiley & Sons, 1970.
[10] F. Oberhettinger, *Tables of Mellin Transform*. Springer, 1974.
[11] J.-M. Nicolas, “Filtrage homomorphique optimal RSO (Radar à Synthèse d’Ouverture),” in *GRETSI’01*, Toulouse, France, Sep. 2001.
[12] W. C. Hoffmann, *Statistical Methods in Radio Wave Propagation*. Pergamon Press, 1960.
[13] H. Cramér, *Mathematical Methods of Statistics*. Princeton University Press, 1946.
[14] J.-M. Nicolas, “Introduction aux statistique de deuxième espèce: Applications aux lois d’images RSO,” ENST, Tech. Rep. 2002D001, 2002.
[15] Z. A. Lomnicki, “On the distribution of products of random variables,” *J. Royal Statistics Soc. Ser. B*, vol. 29, no. 3, pp. 513–524, 1967.
[16] C.-J. Oliver, “Optimum texture estimators for SAR clutter,” *J. Phys. D.: Appl. Phys.*, vol. 26, pp. 1824–1835, 1993.
[17] R.-D. Pierce, “Application of the positive alpha-stable distribution,” in *IEEE Signal Processing Workshop on Higher-Order Statistics*. IEEE, 1997, pp. 420–424.
[18] E. Kuruoğlu and J. Zerubia, “Modelling SAR images with a generalisation of the Rayleigh distribution,” INRIA, Tech. Rep. 4121, 2001.
[19] A. Stuart and J. Keith, Eds., *Kendall’s Advanced Theory of Statistics, Vol. 1: Distribution Theory*, 5th ed. Griffin, 1987.
[20] M.-J. Wasilewski, “Sur certaines propriétés de la distribution Gamma généralisée,” *Revue de Statistique Apliquées*, vol. 15, no. 1, pp. 95–105, 1967.
[21] A. Tagliani, “Numerical inversion of the Mellin transform on the real line for heavy-tailed probability density functions,” *Appl. Math. and Computation*, vol. 130, no. 2-3, pp. 525–536, 2002.
[22] ——, “Recovering a probability density function from its Mellin transform,” *Appl. Math. and Computation*, vol. 118, no. 2-3, pp. 151–159, 2001.
[23] J.-M. Nicolas and A. Maruani, “Numerical Mellin transform applied to texture classification on SAR images,” in *PIERS’99*, Taiwan, Mar. 1999.