Log-Normal approximation of Chi-square distributions for signal processing

Wassim Jouini, Daniel Le Guennec, Christophe Moy, and Jacques Palicot

SUPELEC, SCEE/ETR, Avenue de la Boulaie, CS 47601, 35576 Cesson Sévigné Cedex, France.
Email: \{wassim.jouini, daniel.leguennec, christophe.moy, jacques.palicot\}@supelec.fr

Abstract

We investigate in this paper a Log-Normal approximation of $\chi^2$ distributions. Our study is motivated by the analysis of ratios of random variables where both, Log-Normal and $\chi^2$ distributions are involved. Such ratios appear, for instance, when dealing with an energy detector while only an imperfect knowledge of the noise level is available. In order to characterize the distribution of the ratio, an accurate approximation of the $\chi^2$ distribution by a Log-Normal distribution would highly simplify this problem known to be analytically intractable otherwise.

1 Introduction and Motivations

Let us introduce and motivate the herein developed work based on a case-study on the energy detector and its detection limits under noise uncertainty. The Neyman-Pearson Energy Detector (NP-ED, also known as Energy Detector or radiometric detector) is a commonly used spectrum sensor[1]. The main detection process relies on the comparison of the perceived energy, usually models as $\chi^2$ random variables, to a fixed threshold that depends on the desired performances of the detector as well as the noise power level. When the noise power is only know through an estimation distribution, a classic approach to alleviate this lack of knowledge relies on the analysis of the resulting ratio distribution[2]. It is usually accepted that noise power estimation can be accurately described using Log-Normal distributions. Hence, analyzing the resulting ratio distribution involves the knowledge of the Probability Density Function (PDF) of the ratio statistic composed of a $\chi^2$ and a Log-Normal distributions, which has no simple known forms.

In order to be able to tackle these problems, we suggest to approximate the considered $\chi^2$ distribution by an adequate Log-Normal distribution. As a matter of fact, this approximation reduces the problem to the analysis of a ratio of Log-Normal distributions. Thus in this paper, we investigate, analyze and evaluate a Log-Normal approximation of $\chi^2$ distributions. We show, relying a mild approximations, that Log-Normal distributions offer a good alternative substitute to $\chi^2$ distributions. Moreover, we show the errors of approximation due to a Log-Normal model are smaller than those resulting from the usually suggested Normal distributions.

The rest of this paper is organized as follows : Section2:Model introduces the general model considered. Then, Section 3, develops the mains results and illustrates them with simulation curves. Finally Section 4 concludes.

2 Mathematical Model

Definition 1 (Distributions) Let \( f_{\chi^2_M}(\cdot) \), \( f_{\log N(\mu_L, \sigma^2_L)}(\cdot) \)\(^1\) and \( f_{N(\mu_N, \sigma^2_N)}(\cdot) \) denote, respectively, the PDFs of a $\chi^2$ distribution with \( M \) degrees of freedom, a Log-Normal distribution with parameters \( \{\mu_L, \sigma^2_L\} \) and a normal distribution with parameters \( \{\mu_N, \sigma^2_N\} \), such that:

\(^1\)We use, in this paper, the natural logarithm function \( \log(\cdot) \)
Definition 3 (Approximation and evaluated functions)

Let

\[
\begin{aligned}
f_{\chi_M^2}(x) &= \frac{1}{2^{M/2}\Gamma(M/2)}x^{M/2-1}e^{-x/2}, x \in \mathbb{R}_+, 0 \text{ otherwise} \\
f_{\text{LogN}(\mu, \sigma^2_N)}(x) &= \frac{1}{\sqrt{2\pi \sigma_N^2}}e^{-\frac{(\log(x)-\mu)^2}{2\sigma_N^2}}, x \in \mathbb{R}_+, 0 \text{ otherwise}
\end{aligned}
\]

We consider in the rest of this paper the following specific parameters such that:

\[
\begin{aligned}
\{\mu_L, \sigma_L^2\} &= \{\log(M) - \sigma_L^2/2, \log(1+2/M)\} \\
\{\mu_N, \sigma_N^2\} &= \{M, 2M\}
\end{aligned}
\]

These parameters were chosen such that all three distributions have the same mean and variance.

Definition 2 (Partial Taylor polynomial) Let \(f_D(\cdot)\) the PDF of a distribution \(D \in \{\chi^2, \text{LogN}\}\). We denote by \(T_{D,x_0}(\cdot)\) the following polynomial evaluated at the finite real point \(x_0\):

\[
T_{x_0,D}^{(n)}(x) = \frac{f_D(x)}{f_{N(\mu_v,\sigma^2_v)}(x)} = 1 + \sum_{j=0}^{n} C_{j,D}(x-x_0)^j + \epsilon_D^{(n)}(x)
\]

where \(n\) is the approximation order, \(\{C_{j,D}\}_{j=0,...,n}\) are the polynomial components of the power series and \(\epsilon_D^{(n)}(\cdot)\) is an implicit function that contains the missing terms to respect the equality. \(\epsilon_D^{(n)}(\cdot)\) is very small compared to the other terms that converges to 0 as \(x\) tends to \(x_0\).

In the case of our analysis, the polynomial components of the considered series: \(T_{M,\chi_M^2}^{(n)}(\cdot)\) and \(T_{M,\text{LogN}(\mu,\sigma^2_N)}^{(n)}(\cdot)\) are regular functions of the parameter \(M\). Their expression is usually very complex. However, since we are only interested in the asymptotic behavior of these functions, we simplify the general expression of these parameters using the first existing order of their polynomial expression evaluated as \(M\) tends to infinity such that for all \(j = \{1, \cdots, n\}:

\[
C_{j,D}(k) = \sum_{i=0}^{\infty} \frac{c_{i,j,D}}{k^i} \approx \frac{c_{i_0,j,D}}{k^{i_0}}
\]

where \(i_0(j)\) is the first index in \(\mathbb{N}\) such that \(c_{i_0(j),j,D} \neq 0\).

Definition 3 (Approximation and evaluated functions) Let \(\tilde{C}_{j,D}(k) = \frac{c_{i_0(j),j,D}}{k^{i_0}}\) be the asymptotically approximated polynomial component, and \(\tilde{T}_{x_0,D}^{(n)}(\cdot)\) the partial approximation of Taylor series as defined in Definition 2:

\[
\tilde{T}_{x_0,D}^{(n)}(x) = 1 + \sum_{j=0}^{n} \tilde{C}_{j,D}(x-x_0)^j
\]

We evaluate in this paper the asymptotic behavior of the following error functions for large \(M\):

\[
\begin{aligned}
\Delta_1(x) &= f_{\text{LogN}(\mu,\sigma^2_N)}(x) - f_{\chi_M^2}(x) \approx \left(\tilde{T}_{M,\text{LogN}(\mu,\sigma^2_N)}^{(n)}(x) - \tilde{T}_{M,\chi_M^2}^{(n)}(x)\right) f_{N(\mu_v,\sigma^2_v)}(x) \\
\Delta_2(x) &= f_{\chi_M^2}(x) - f_{N(\mu_v,\sigma^2_v)}(x) \approx \left(\tilde{T}_{M,\chi_M^2}^{(n)}(x) - 1\right) f_{N(\mu_v,\sigma^2_v)}(x)
\end{aligned}
\]

In the rest of this section, we focus on the analysis of the approximations :

\[
\begin{aligned}
\tilde{\Delta}_1(x) &= \left(\tilde{T}_{M,\text{LogN}(\mu,\sigma^2_N)}^{(n)}(x) - \tilde{T}_{M,\chi_M^2}^{(n)}(x)\right) f_{N(\mu_v,\sigma^2_v)}(x) \\
\tilde{\Delta}_2(x) &= \left(\tilde{T}_{M,\chi_M^2}^{(n)}(x) - 1\right) f_{N(\mu_v,\sigma^2_v)}(x)
\end{aligned}
\]
Figure 1: Approximation Error functions. Both left and right figures plot the functions $\Delta_1(\cdot)$, $\Delta_2(\cdot)$, $\tilde{\Delta}_1(\cdot)$ and $\tilde{\Delta}_2(\cdot)$. However, the left figure shows these functions for a parameter $M = 1000$, while the right figure shows them for $M = 25000$. We observe in this figure that the theoretical approximation introduced in Property 1 seems to converge to the real values as $M$ grows large.

3 Main Results

We focus on an approximation of the third order $n = 3$ and analyze the extrema of the functions $\tilde{\Delta}_1(\cdot)$ and $\tilde{\Delta}_2(\cdot)$. The choice of a third order approximation was motivated by the necessity of obtaining analytical solutions for the extrema.

Property 1 (Approximated error functions) Let the function $\tilde{\Delta}_1(\cdot)$ and $\tilde{\Delta}_2(\cdot)$ two approximation errors function as defined in Equation 7, then we can show that:

$$
\begin{align*}
\tilde{\Delta}_1(x) &= \left(-\frac{M}{6M^2} - \frac{M}{2M^2} + \frac{(x-M)^2}{6M^2} + \frac{(x-M)^3}{8M^2}\right)f_N(\mu, \sigma^2)(x) \\
\tilde{\Delta}_2(x) &= \left(-\frac{5}{12M} - \frac{M}{2M^2} + \frac{5(x-M)^2}{8M^2} + \frac{(x-M)^3}{12M^2}\right)f_N(\mu, \sigma^2)(x)
\end{align*}
$$

Property 2 (Extrema : position and amplitude) Let the real values $\{y_{1,i}(M)\}_{i=1}^4$ and $\{y_{2,i}(M)\}_{i=1}^4$ denote the approximated extrema amplitudes, of, respectively, $\tilde{\Delta}_1(\cdot)$ and $\tilde{\Delta}_2(\cdot)$ at the positions $\{x_1(M) > x_2(M) > x_3(M) > x_4(M)\}$. Then there exist two real constants $\{a, b\}$ such that: $\{a, b\} = \{4 - 2^\frac{1}{4} - 2^\frac{3}{4}, 2 + 2^\frac{1}{4} + 2^\frac{3}{4}\}$ such that for large $M$:

$$
\begin{align*}
x_1(M) &\approx M + \left(-\sqrt{a} - \sqrt{b}\right)\sqrt{M} ; \ x_2(M) \approx M + \left(\sqrt{a} - \sqrt{b}\right)\sqrt{M} \\
x_3(M) &\approx M + \left(-\sqrt{a} + \sqrt{b}\right)\sqrt{M} ; \ x_4(M) \approx M + \left(\sqrt{a} + \sqrt{b}\right)\sqrt{M}
\end{align*}
$$

Which leads to the following expressions for the approximated extrema of $\tilde{\Delta}_1(\cdot)$ for large $M$:

$$
\begin{align*}
\tilde{\Delta}_1(x_1(M)) &\approx y_{1,1}(M) \approx \frac{e^{-\frac{1}{4}(\sqrt{a} + \sqrt{b})^2}}{24\sqrt{\pi}} \left(-\frac{2}{3} \sqrt{ab} \left(\sqrt{a} + \sqrt{b}\right) + \frac{5}{2M^{3/2}} \left(2 + 3(\sqrt{a} + \sqrt{b})^2\right)\right) \\
\tilde{\Delta}_1(x_2(M)) &\approx y_{1,2}(M) \approx \frac{e^{-\frac{1}{4}(\sqrt{a} - \sqrt{b})^2}}{24\sqrt{\pi}} \left(-\frac{2}{3} \sqrt{ab} \left(\sqrt{a} - \sqrt{b}\right) + \frac{5}{2M^{3/2}} \left(38 - 3(\sqrt{a} + \sqrt{b})^2\right)\right) \\
\tilde{\Delta}_1(x_3(M)) &\approx y_{1,3}(k) \approx \frac{e^{-\frac{1}{4}(\sqrt{a} + \sqrt{b})^2}}{24\sqrt{\pi}} \left(\frac{2}{3} \sqrt{ab} \left(\sqrt{a} - \sqrt{b}\right) + \frac{5}{2M^{3/2}} \left(38 - 3(\sqrt{a} + \sqrt{b})^2\right)\right) \\
\tilde{\Delta}_1(x_4(M)) &\approx y_{1,4}(M) \approx \frac{e^{-\frac{1}{4}(\sqrt{a} - \sqrt{b})^2}}{24\sqrt{\pi}} \left(\frac{2}{3} \sqrt{ab} \left(\sqrt{a} + \sqrt{b}\right) + \frac{5}{2M^{3/2}} \left(2 + 3(\sqrt{a} + \sqrt{b})^2\right)\right)
\end{align*}
$$
Figure 2: Maximum Absolute Error. In this figure, four curves are represented: two of them, in solid line, illustrate the decreasing rate of the global maximum of the error functions $\Delta_1(\cdot)$ and $\Delta_2(\cdot)$. Whereas, the two other curves, plot the theoretical maximum of the approximations $\tilde{\Delta}_1(\cdot)$ and $\tilde{\Delta}_2(\cdot)$.

As well as the following expressions for the approximated extrema of $\tilde{\Delta}_2(\cdot)$ for large $k$:

\[
\begin{align*}
\tilde{\Delta}_2(x_1(M)) & \approx y_{2,1}(M) \approx e^{-\frac{1}{2}\left(\frac{5\pi+\sqrt{5}}{2}\right)^2} \left( -\frac{2}{M^{3/2}} \sqrt{ab} \left( \sqrt{a} + \sqrt{b} \right) + \frac{1}{M^{3/2}} \left( -1 + 6(\sqrt{a} + \sqrt{b})^2 \right) \right) \\
\tilde{\Delta}_2(x_2(M)) & \approx y_{2,2}(M) \approx e^{-\frac{1}{2}\left(\frac{71}{12}\pi + \sqrt{5}\right)^2} \left( -\frac{2}{M^{3/2}} \sqrt{ab} \left( \sqrt{a} - \sqrt{b} \right) + \frac{1}{M^{3/2}} \left( 71 - 6(\sqrt{a} + \sqrt{b})^2 \right) \right) \\
\tilde{\Delta}_2(x_3(M)) & \approx y_{2,3}(M) \approx e^{-\frac{1}{2}\left(\frac{71}{12}\pi - \sqrt{5}\right)^2} \left( \frac{2}{M^{3/2}} \sqrt{ab} \left( \sqrt{a} - \sqrt{b} \right) + \frac{1}{M^{3/2}} \left( 71 - 6(\sqrt{a} + \sqrt{b})^2 \right) \right) \\
\tilde{\Delta}_2(x_4(M)) & \approx y_{2,4}(M) \approx e^{-\frac{1}{2}\left(\frac{71}{12}\pi + \sqrt{5}\right)^2} \left( \frac{2}{M^{3/2}} \sqrt{ab} \left( \sqrt{a} + \sqrt{b} \right) + \frac{1}{M^{3/2}} \left( -1 + 6(\sqrt{a} + \sqrt{b})^2 \right) \right)
\end{align*}
\]

Here is a sketch of the followed mathematical protocol to obtain the stated results. Let us consider the approximation functions $\Delta_1(\cdot)$ and $\Delta_2(\cdot)$ as defined in Equation 8. The analysis of their derivative functions is equivalent to a root analysis of a fourth degree polynomial, which can be solved using the well known Ferrari approach. This latter provides us with values which leading terms (for large $M$) are equal to $\left\{x_i(M)\right\}_{i=1}^4$. These solutions appear to be the same for both error functions. Equations 10 and 11 are, then, computed as the evaluation of $\Delta_1(\cdot)$ and $\Delta_2(\cdot)$ at the positions $\left\{x_i(M)\right\}_{i=1}^4$.

These results supported by the empirical illustrations in Figure 2 and 1 suggest that the Log-Normal approximation offers a satisfactory approximations for $\chi^2$ distributions, opening the field to many possible signal processing applications.

### 4 Conclusion

We investigated in this paper, a Log-Normal approximation of $\chi^2$ distributions. Relying on both, theoretical and empirical evaluations, we showed that Log-Normal distributions offers a satisfactory approximations for $\chi^2$ distributions. Moreover, we showed that the error due to the approximation, is smaller in the case of Log-Normal approximation compared to the usually suggested Normal approximation.

### References
