Université de Montréal

# Quadratic Distance Methods Applied to Generalized Normal Laplace Distribution

par

Ionica Groparu-Cojocaru

Département de mathématiques et de statistique
Faculté des arts et des sciences

Mémoire présenté à la Faculté des études supérieures
en vue de l'obtention du grade de
Maître ès sciences (M.Sc.)
en mathématiques

Orientation actuariat

mars 2007

# Université de Montréal

Faculté des études supérieures

Ce mémoire intitulé

## Quadratic Distance Methods Applied to Generalized Normal Laplace Distribution

présenté par


## Ionica Groparu-Cojocaru


a été évalué par un jury composé des personnes suivantes :

*Martin Goldstein*

(président-rapporteur)

*Louis G. Doray*

(directeur de recherche)

*Charles Dugas*

(membre du jury)


Mémoire accepté le:

04/04/07

# SOMMAIRE

Le but de cette thèse est d'analyser la méthode d'estimation des paramètres de la distribution normale généralisée de Laplace (GNL).
Nous proposons une méthode alternative pour estimer les paramètres basée sur la minimisation de la distance quadratique entre les parties réelle et imaginaire des fonctions caractéristiques empiriques et théoriques.
L'estimateur de distance quadratique (QDE) obtenu est convergent, robuste, asymptotiquement sans biais et de distribution normale. On développe des statistiques pour les tests d'ajustement et nous démontrons que les tests suivent asymptotiquement une distribution chi-carré.
Les résultats de la simulation sont fournis et les propriétés asymptotiques des estimateurs obtenus sont étudiées.

# SUMMARY

The purpose of this thesis is to analyze the method of parameter estimation of the generalized normal Laplace distribution (GNL).

We propose an alternative method for estimating the parameters based on minimizing the quadratic distance between the real and imaginary parts of empirical and theoretical characteristic functions.

We show that the quadratic distance estimator (QDE) obtained is consistent, robust, asymptotically unbiased and normally distributed. We develop test statistics for goodness-of-fit and prove that these tests follow asymptotically a chi-square distribution.

Simulation results are provided and asymptotic properties of the estimators obtained are studied.

Keywords: Empirical characteristic function, normal Laplace distribution, generalized normal Laplace distribution, quadratic distance estimator, influence function, robustness, goodness-of-fit tests, simple and composite hypotheses.

# CONTENTS

# LIST OF TABLES

# ACKNOWLEDGEMENTS

I express my entire gratitude to my supervisor Prof. Louis G. Doray who offered an insightful guidance towards the progress of my project and who supervised this work with professionalism, interest and steady encouragements.

I acknowledge the ongoing support received from Prof. Roch Roy and Prof. Marlène Frigon. I am also grateful to the professors from the Department of Mathematics and Statistics for their scientific and valuable lectures.

I greatly benefited from the help provided by the informatics counselor Francis Forget and I would like to thank him.

For financial support, I would like to acknowledge FQRNT, the Mathematics and Statistics Department, Fondation J.A. Bombardier and my supervisor Prof. Louis Doray.

# INTRODUCTION

Although the normal (Gaussian) distribution plays a central role in basic statistics, it has long been recognized that the empirical distributions of many phenomena modelled by the normal distribution sometimes do not follow closely to the Gaussian shape.

In recent years the huge burst of research interest in financial modelling along with the availability of high frequency price data and the concurrent understanding that logarithmic price returns do not follow exactly a normal distribution (see e.g. Rydberg, 2000), as previously assumed, has led to a search for more realistic alternative parametric models.

Reed (2004) introduced new distributions namely, normal-Laplace, generalized normal Laplace and double Pareto-lognormal and proved their usefulness in modelling the size distributions of various phenomena arising in a wide range of areas of inquiry such as economics, finance, geography, physical sciences and geology.

In this thesis, we focus on the properties and estimation procedure for the generalized normal Laplace (GNL) distribution. To date, attempts to estimate the parameters of GNL distribution have been made with the maximum likelihood method and the method of moments. The lack of a closed form for the GNL density generates difficulties for estimation by maximum likelihood. Even though the method of moments estimates are consistent and asymptotically normal, they are not generally efficient (not achieving the Cramér-Rao bound) even asymptotically.

Based on these facts, we propose the quadratic distance method as an alternative method to the ones mentioned above. Luong and Thompson (1987) introduced estimators based on a "quadratic transform distance" and derived the asymptotic properties of these estimators.

The thesis consists of five chapters and we provide their description in the following.

In chapter 1 the generalized normal Laplace distribution is defined and its properties are proved. The normal Laplace and the double Pareto-lognormal distributions are presented with a list of some of their properties. Applications of these distributions in financial modelling are discussed.

In chapter 2 we present the quadratic transform distance estimators. The influence function is derived and sufficient conditions for their consistency and asymptotic normality are given. Corresponding goodness-of-fit tests in cases of simple and composite null hypotheses are developed.

In chapter 3 we develop the quadratic distance estimators based on the characteristic function in order to estimate the parameters of the generalized normal Laplace distribution. This is a special distance within the class of quadratic transform distances presented in chapter 2. We develop an expression for the variance-covariance matrix of the errors between the empirical and characteristic functions. Properties such as robustness, consistency and asymptotic normality are established. We present goodness-of-fit test statistics and we show that they have an asymptotic chi-square distribution.

Numerical simulation results are provided in chapter 4 and they indicate the performance of these estimators. FindMinimum procedure in MATHEMATICA is used, which proved to be very quick in this case.

We draw conclusions in chapter 5.

# Chapter 1

## THE GENERALIZED NORMAL LAPLACE DISTRIBUTION

The most effective tool in the study of probability distributions is the use of their functional transforms. The moment generating function and the characteristic function are defined in the first section of this chapter. Properties of these transforms such as the multiplication property, that is summation of independent random variables corresponds to multiplication of their transforms and the uniqueness, that is if two random variables have the same transform then they also have the same distribution, are proved.

The normal Laplace (NL) distribution and the double Pareto-lognormal distribution are introduced and some of their properties are mentioned.

Reed (2004) introduced a new infinitely divisible distribution, namely the generalized normal Laplace (GNL) distribution, which exhibits the properties seen in observed logarithmic returns. The (GNL) distribution is an extension of the (NL) distribution. Properties of the GNL distribution proved in this chapter include its infinitely divisibility, its moments and its closure under the convolution operation, i.e. sums of independent GNL random variables follow GNL distributions.

Applications of these distributions in financial modelling are presented.

The maximum likelihood estimation method and the method of moments for the generalized normal Laplace distribution are discussed in this chapter.

## 1.1. Moment generating function and characteristic function

Some of the most important tools in probability theory are borrowed from other branches of mathematics. In this section we discuss two such closely related tools. We begin with moment generating functions and then treat characteristic functions.

**Definition 1.1.1.** *Let $\mathbf{X}$ be a random variable. The moment generating function of $\mathbf{X}$ is defined by the following formula, where "$E$" means expected value:*

$$M_{\mathbf{X}}(t) = E[\exp(t\mathbf{X})]$$

*provided the expectation is finite for $|t| < h$ for some $h > 0$.*
*If $F_{\mathbf{X}}(x)$ is the cumulative distribution function of $\mathbf{X}$, then the moment generating function is given by the Riemann-Stieltjes integral:*

$$M_{\mathbf{X}}(t) = E[\exp(t\mathbf{X})] = \int_{-\infty}^{\infty} \exp(tx)dF_{\mathbf{X}}(x).$$

*In cases in which there is a probability density function, $f_{\mathbf{X}}(x)$, this becomes*

$$M_{\mathbf{X}}(t) = E[\exp(t\mathbf{X})] = \int_{-\infty}^{\infty} \exp(tx)f_{\mathbf{X}}(x)dx.$$

We note that $M_{\mathbf{X}}(0) = 1$.

**Theorem 1.1.1.** *Let $\mathbf{X}$ be a random variable whose moment generating function, $M_{\mathbf{X}}(t)$, exists for $|t| < h$ for some $h > 0$. Then*
*a) $E[|\mathbf{X}|^r] < \infty$ for all $r > 0$;*
*b) $E[\mathbf{X}^n] = \frac{d^n}{dt^n}M_{\mathbf{X}}(t)|_{t=0} = M_{\mathbf{X}}^{(n)}(0)$ for $n = 1, 2, \ldots$ .*

Proof. *a)* Let $r > 0$ and $|t| < h$ be given. Since $|x|^r/\exp(|tx|) \to 0$ as $x \to \infty$ for all $r > 0$, we choose $A$ (depending on $r$) in such way that

$$|x|^r \leq \exp(|tx|) \quad \textit{whenever} \quad |x| > A.$$

Then

$$E[|\mathbf{X}|^r] = E[|\mathbf{X}|^r]I(|\mathbf{X}| \leq A) + E[|\mathbf{X}|^r]I(|\mathbf{X}| > A)$$

$$\leq A^r + E[\exp(|t\mathbf{X}|)]I(|\mathbf{X}| > A) \leq A^r + E[\exp(|t\mathbf{X}|)] < \infty$$

since $\exp(|tx|) \leq \exp(tx) + \exp(-tx)$ and $M_{\mathbf{X}}(t) = E[\exp(t\mathbf{X})]$ exists for $|t| < h$ for some $h > 0$.

$b)$ Using the previous result, we obtain by differentiation (under the integral sign):

$$M_{\mathbf{X}}^{(n)}(t) = \int_{-\infty}^{\infty} x^n \exp(tx) dF_{\mathbf{X}}(x)$$

which yields

$$M_{\mathbf{X}}^{(n)}(0) = \int_{-\infty}^{\infty} x^n dF_{\mathbf{X}}(x) = E[\mathbf{X}^n].$$

□

Following are two examples listing the moment generating functions of the normal and gamma distributions.

***Example 1.*** Let $\mathbf{X}$ be a random variable normally distributed with mean $\mu$ and variance $\sigma^2$, that is $\mathbf{X} \sim \mathbf{N}(\mu, \sigma^2)$ with the probability density function (pdf) given by

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right), \quad x \in \mathbf{R}.$$

Then

$$M_{\mathbf{X}}(t) = E[\exp(t\mathbf{X})] = \int_{-\infty}^{\infty} \exp(tx) \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-(x-\mu)^2/2\sigma^2\right) dx$$

$$= \int_{-\infty}^{\infty} \exp(t(y+\mu)) \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-y^2/2\sigma^2\right) dy$$

$$= \exp(\mu t) \int_{-\infty}^{\infty} \frac{1}{\sigma\sqrt{2\pi}} \exp\left(ty - y^2/2\sigma^2\right) dy.$$

Now

$$ty - y^2/2\sigma^2 = -(y - \sigma^2 t)^2/2\sigma^2 + \sigma^2 t^2/2.$$

Consequently,

$$M_{\mathbf{X}}(t) = \exp(\mu t)\exp(\sigma^2 t^2/2) \int_{-\infty}^{\infty} \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-(y-\sigma^2 t)^2/2\sigma^2\right) dy.$$

Since the last integral represents the integral of the normal density $\mathbf{N}(\sigma^2 t,\, \sigma^2)$, its value is one and therefore

$$M_{\mathbf{X}}(t) = \exp\left(\mu t + \sigma^2 t^2/2\right), \quad -\infty < t < \infty.$$

***Example 2.*** Let $\mathbf{X}$ be a random variable which follows a gamma distribution with shape parameter $\rho > 0$ and scale parameter $\theta > 0$ that is, $\mathbf{X} \sim \mathbf{Gamma}(\rho,\, \theta)$ with the probability density function (pdf) given by

$$f(x) = \frac{\theta^\rho}{\Gamma(\rho)} x^{\rho-1} \exp(-\theta x), \quad x > 0,$$

where $\Gamma$ is the gamma function defined by

$$\Gamma(\alpha) = \int_0^{\infty} x^{\alpha-1} \exp(-x) dx, \quad \alpha > 0,$$

then

$$M_{\mathbf{X}}(t) = \int_0^{\infty} \exp(tx) \frac{\theta^\rho}{\Gamma(\rho)} x^{\rho-1} \exp(-\theta x) dx$$

$$= \frac{\theta^\rho}{\Gamma(\rho)} \int_0^{\infty} x^{\rho-1} \exp\left(-(\theta - t)x\right) dx$$

$$= \frac{\theta^\rho}{\Gamma(\rho)} \frac{\Gamma(\rho)}{(\theta - t)^\rho}$$

for $-\infty < t < \theta$. The integral diverges for $\theta \leq t < \infty$. Thus, for $-\infty < t < \theta$ we have

$$M_{\mathbf{X}}(t) = (1 - t/\theta)^{-\rho}.$$

The following definition introduces the characteristic function of a random variable.

**Definition 1.1.2.** *The characteristic function of any probability distribution on the real line is defined by the following formula, where* $\mathbf{X}$ *is a random variable with the distribution in question:*

$$\phi_{\mathbf{X}}(t) = E[\exp(it\mathbf{X})], \quad i^2 = -1$$

*where "E" means expected value, t is a real number and* $\exp(itx)$ *is the function for real x defined by*

$$\exp(itx) = \cos(tx) + i\sin(tx).$$

*If* $F_{\mathbf{X}}(x)$ *is the cumulative distribution function of* $\mathbf{X}$*, then the characteristic function is given by the Riemann-Stieltjes integral:*

$$\phi_{\mathbf{X}}(t) = E[\exp(it\mathbf{X})] = \int_{-\infty}^{\infty} \exp(itx)dF_{\mathbf{X}}(x).$$

*In cases in which there is a probability density function,* $f_{\mathbf{X}}(x)$*, this becomes*

$$\phi_{\mathbf{X}}(t) = E[\exp(it\mathbf{X})] = \int_{-\infty}^{\infty} \exp(itx)f_{\mathbf{X}}(x)dx.$$

We note that $\phi_{\mathbf{X}}(0) = 1$.

The following theorem gives some basic properties of the characteristic functions. The first property tells us that characteristic functions exist for all random variables.

**Theorem 1.1.2.** *Let* $\mathbf{X}$ *be a random variable with the characteristic function* $\phi_{\mathbf{X}}(t)$*. Then*

*a)* $|\phi_{\mathbf{X}}(t)| \leq \phi_{\mathbf{X}}(0) = 1$ *for all real values of* $t$*;*

*b)* $\overline{\phi_{\mathbf{X}}(t)} = \phi_{\mathbf{X}}(-t) = \phi_{-\mathbf{X}}(t)$ *for all real* $t$*, where* $\overline{\phi_{\mathbf{X}}(t)}$ *is the complex conjugate of* $\phi_{\mathbf{X}}(t)$*;*

*c)* $\phi_{\mathbf{X}}(t)$ *is uniformly continuous;*

*d)* $a\mathbf{X} + b$ *has the characteristic function*

$$\phi_{a\mathbf{X}+b}(t) = \exp(ibt)\phi_{\mathbf{X}}(at), \quad t \in \mathbf{R}.$$

PROOF. *a*) Because $|\exp(itx)| = 1$ for all $t$ and $x$ we obtain

$$|\phi_{\mathbf{X}}(t)| = |E[\exp(it\mathbf{X})]| \le \int_{-\infty}^{\infty} |\exp(itx)| dF_{\mathbf{X}}(x) = \int_{-\infty}^{\infty} dF_{\mathbf{X}}(x) = 1 = \phi_{\mathbf{X}}(0).$$

*b*) Using the fact that $\cos x$ is even and $\sin x$ is odd, that is $\cos(-x) = \cos x$ and $\sin(-x) = -\sin x$ for all $x$, we prove this result as follows:

$$\overline{\exp(ixt)} = \overline{\cos(xt) + i\sin(xt)} = \cos(xt) - i\sin(xt)$$

$$= \cos(x(-t)) + i\sin(x(-t)) = \exp(ix(-t))$$

$$= \cos((-x)t) + i\sin((-x)t) = \exp(i(-x)t).$$

Therefore, $\overline{\phi_{\mathbf{X}}(t)} = \phi_{\mathbf{X}}(-t) = \phi_{-\mathbf{X}}(t)$.

*c*) Let $t$ be arbitrary and $h > 0$ (a similar argument works for $h < 0$).

Before we prove the result, we show the following inequalities: $|\exp(ix) - 1| \le 2$ and $|\exp(ix) - 1| \le |x|$ for any real $x$.

For the first inequality we have

$$|\exp(ix) - 1| = |\cos x + i\sin x - 1| = |-2\sin^2(x/2) + 2i\sin(x/2)\cos(x/2)|$$

$$= 2|\sin(x/2)||\sin(x/2) - i\cos(x/2)| \le 2$$

since $|\sin(x/2)| \le 1$ and $|\sin(x/2) - i\cos(x/2)| = 1$ for all real $x$.

For the second inequality we have

$$|\exp(ix) - 1| = |\cos x + i\sin x - 1| = |-2\sin^2(x/2) + 2i\sin(x/2)\cos(x/2)|$$

$$= 2|\sin(x/2)||\sin(x/2) - i\cos(x/2)| \le 2|x/2| = |x|$$

since $|\sin x| \le |x|$ and $|\sin(x/2) - i\cos(x/2)| = 1$ for all real $x$.

For $A > 0$ and using these two inequalities, we obtain

$$|\phi_{\mathbf{X}}(t+h) - \phi_{\mathbf{X}}(t)| = |E[\exp(i(t+h)\mathbf{X})] - E[\exp(it\mathbf{X})]|$$

$$= |E[\exp(it\mathbf{X})(\exp(ih\mathbf{X}) - 1)]| \le E[|\exp(it\mathbf{X})(\exp(ih\mathbf{X}) - 1)|]$$

$$= E[|\exp(ih\mathbf{X}) - 1|] = E[|\exp(ih\mathbf{X}) - 1|I(|\mathbf{X}| \le A)] + E[|\exp(ih\mathbf{X}) - 1|I(|\mathbf{X}| > A)]$$

$$\le E[|h\mathbf{X}|I(|\mathbf{X}| \le A)] + 2P(|\mathbf{X}| > A) \le hA + 2P(|\mathbf{X}| > A) < \epsilon$$

for any $\epsilon > 0$ if we have first chosen $A$ so large that $2P(|\mathbf{X}| > A) < \frac{\epsilon}{2}$ , and then $h$ so small that $hA < \frac{\epsilon}{2}$ . This proves that $\phi_{\mathbf{X}}(t)$ is uniformly continuous.

*d)*

$$\phi_{a\mathbf{X}+b}(t) = E[\exp(i(a\mathbf{X}+b)t] = E[\exp(ia\mathbf{X}t)\exp(ibt)]$$

$$= \exp(ibt)E[\exp(i(at)\mathbf{X})] = \exp(ibt)\phi_{\mathbf{X}}(at).$$

□

We saw earlier that for real values of $t$, the moment generating function, $M(t)$, may not always exist, but for such values the characteristic function, $\phi(t)$, always exists since $|\exp(itx)|$ is a bounded and continuous function for all finite real $t$ and $x$. Similarly $M(t)$ always exists when $t$ is purely imaginary and we have $\phi(t) = M(it)$. Consequently, based on the examples 1 and 2, we can obtain the characteristic functions of the normal and gamma distributions.

If $\mathbf{X} \sim \mathbf{N}(\mu, \sigma^2)$, then its characteristic function has the following expression

$$\phi_{\mathbf{X}}(t) = \exp(\mu i t - \sigma^2 t^2/2), \quad t \in \mathbf{R}. \tag{1.1.1}$$

If $\mathbf{X} \sim \mathbf{Gamma}(\rho, \theta)$ then its characteristic function has the following expression

$$\phi_{\mathbf{X}}(t) = (1 - it/\theta)^{-\rho}, \quad t \in \mathbf{R}. \tag{1.1.2}$$

The multiplication and the uniqueness properties of both characteristic function and moment generating function are established by the following theorems.

**Theorem 1.1.3.** *Let $\mathbf{X}_1, \mathbf{X}_2, ...., \mathbf{X}_n$ be random variables and $h_1, h_2, ...., h_n$ be measurable functions.*

*If $\mathbf{X}_1, \mathbf{X}_2, ...., \mathbf{X}_n$ are independent, then so are $h_1(\mathbf{X}_1), h_2(\mathbf{X}_2), ...., h_n(\mathbf{X}_n)$.*

PROOF. Let $(\Omega, \mathfrak{F}, \mathrm{P})$ be a probability space and $\mathrm{A}_1, \mathrm{A}_2, ...., \mathrm{A}_n$ be Borel measurable sets. Since $\mathbf{X}_1, \mathbf{X}_2, ...., \mathbf{X}_n$ are independent random variables, we have

$$P[\bigcap_{k=1}^{n} h_k(\mathbf{X}_k) \in \mathrm{A}_k] = P[\bigcap_{k=1}^{n} \mathbf{X}_k \in h_k^{-1}(\mathrm{A}_k)]$$

$$= \prod_{k=1}^{n} P[\mathbf{X}_k \in h_k^{-1}(\mathrm{A_k})] = \prod_{k=1}^{n} P[h_k(\mathbf{X}_k) \in \mathrm{A_k}]$$

Thus $h_1(\mathbf{X}_1), h_2(\mathbf{X}_2), ...., h_n(\mathbf{X}_n)$ are independent. □

**Theorem 1.1.4.** *Let $\mathbf{X}_1, \mathbf{X}_2, ...., \mathbf{X}_n$ be independent random variables with the characteristic functions $\phi_{\mathbf{X}_1}, \phi_{\mathbf{X}_2}, ...., \phi_{\mathbf{X}_n}$ respectively and $\mathbf{S}_n = \mathbf{X}_1 + \mathbf{X}_2 + ....\mathbf{X}_n$. Then the characteristic function of $\mathbf{S}_n$ is*

$$\phi_{\mathbf{S}_n}(t) = \phi_{\mathbf{X}_1}(t)\phi_{\mathbf{X}_2}(t)....\phi_{\mathbf{X}_n}(t), \quad t \in \mathbf{R}.$$

*If, in addition, $\mathbf{X}_1, \mathbf{X}_2, ...., \mathbf{X}_n$ are equidistributed, then $\phi_{\mathbf{S}_n}(t) = [\phi_{\mathbf{X}_1}(t)]^n$.*

Proof. Since $\mathbf{X}_1, \mathbf{X}_2, ...., \mathbf{X}_n$ are independent, we know from theorem 1.1.3. that the same is true for $\exp(it\mathbf{X}_1), \exp(it\mathbf{X}_2), ...., \exp(it\mathbf{X}_n)$, so that

$$\phi_{\mathbf{S}_n}(t) = E[\exp(it(\mathbf{X}_1 + \mathbf{X}_2 + .... + \mathbf{X}_n)] = E[\exp(it\mathbf{X}_1)\exp(it\mathbf{X}_2)....\exp(it\mathbf{X}_n)]$$

$$= E[\exp(it\mathbf{X}_1)]E[\exp(it\mathbf{X}_2)]....E[\exp(it\mathbf{X}_n)] = \phi_{\mathbf{X}_1}(t)\phi_{\mathbf{X}_2}(t)....\phi_{\mathbf{X}_n}(t).$$

The second part is immediate, since all factors are the same. □

**Theorem 1.1.5.** *Let $\mathbf{X}_1, \mathbf{X}_2, ...., \mathbf{X}_n$ be independent random variables whose moment generating functions exist for $|t| < h$ for some $h > 0$, and let us define $\mathbf{S}_n = \mathbf{X}_1 + \mathbf{X}_2 + ....\mathbf{X}_n$. Then*

$$M_{\mathbf{S}_n}(t) = M_{\mathbf{X}_1}(t)M_{\mathbf{X}_2}(t)....M_{\mathbf{X}_n}(t), \ |t| < h.$$

*If in addition, $\mathbf{X}_1, \mathbf{X}_2, ...., \mathbf{X}_n$ are equidistributed, then*

$$M_{\mathbf{S}_n}(t) = [M_{\mathbf{X}_1}(t)]^n, \ |t| < h.$$

Proof. Since $\mathbf{X}_1, \mathbf{X}_2, ...., \mathbf{X}_n$ are independent, we know from theorem 1.1.3. that the same is true for $\exp(t\mathbf{X}_1), \exp(t\mathbf{X}_2), ...., \exp(t\mathbf{X}_n)$, so that

$$M_{\mathbf{S}_n}(t) = E[\exp(t(\mathbf{X}_1 + \mathbf{X}_2 + .... + \mathbf{X}_n)]$$

$$= E[\exp(t\mathbf{X}_1)\exp(t\mathbf{X}_2)....\exp(t\mathbf{X}_n)]$$

$$= E[\exp(t\mathbf{X}_1)]E[\exp(t\mathbf{X}_2)]....E[\exp(t\mathbf{X}_n)]$$

$$= M_{\mathbf{X}_1}(t) M_{\mathbf{X}_2}(t) .... M_{\mathbf{X}_n}(t).$$

The second part is immediate, since all factors are the same. □

**Theorem 1.1.6.** *Let* $\mathbf{X}$ *and* $\mathbf{Y}$ *be random variables with the characteristic functions* $\phi_{\mathbf{X}}$ *and* $\phi_{\mathbf{Y}}$ *respectively.*
*If* $\phi_{\mathbf{X}} = \phi_{\mathbf{Y}}$, *then* $\mathbf{X} \stackrel{d}{=} \mathbf{Y}$ *and conversely (the equality holds in distribution and conversely).*

PROOF. See Shiryayev (1984). □

**Theorem 1.1.7.** *Let* $\mathbf{X}$ *and* $\mathbf{Y}$ *be random variables with the moment generating functions* $M_{\mathbf{X}}(t)$ *and* $M_{\mathbf{Y}}(t)$ *respectively.*
*If* $M_{\mathbf{X}}(t) = M_{\mathbf{Y}}(t)$ *where* $|t| < h$ *for some* $h > 0$, *then* $\mathbf{X} \stackrel{d}{=} \mathbf{Y}$.

PROOF. A moment generating function which, as required, is finite in the interval $|t| < h$ can, according to results from the theory of analytic functions, be extended to a complex function $E[\exp(z\mathbf{X})]$ for $|\mathrm{Re}(z)| < h$. Putting $z = iy$, where $y$ is real, yields the characteristic function.

Thus, if two moment generating functions are equal, then so are the corresponding characteristic functions, which uniquely determines the distribution according to theorem 1.1.6.. □

## 1.2. Genesis, properties and financial applications of the NL and GNL distributions

In this section, the generalized normal Laplace distribution (GNL) is defined and its representation as a sum of independent normal and gamma random variables is proved. The GNL distribution is an extension of the normal Laplace (NL) distribution.

We begin by presenting the NL distribution and the double Pareto-lognormal distribution which is related to the NL distribution.

**Definition 1.2.1.** *The normal Laplace (NL) distribution can be defined in terms of its cumulative distribution function (cdf) which for all real $x$ is*

$$F(x) = \Phi(\frac{x-\mu}{\sigma}) - \phi(\frac{x-\mu}{\sigma})\frac{\beta R(\alpha\sigma - (x-\mu)/\sigma) - \alpha R(\beta\sigma + (x-\mu)/\sigma)}{\alpha+\beta},$$

*where $\Phi$ and $\phi$ are the cdf and probability density function (pdf) of a standard normal random variable and $R$ is Mills' ratio:*

$$R(z) = \frac{\Phi^c(z)}{\phi(z)} = \frac{1-\Phi(z)}{\phi(z)} = \frac{1}{h(z)},$$

*where $h(z)$ is the hazard rate.*

*The location parameter $\mu$ can assume any real value while the scale parameter $\sigma$ and the other two parameters $\alpha$ and $\beta$, which determine tail behavior, are assumed to be positive.*

*The corresponding density (pdf) is*

$$f(x) = \frac{\alpha\beta}{\alpha+\beta}\phi(\frac{x-\mu}{\sigma})[R(\alpha\sigma - (x-\mu)/\sigma) + R(\beta\sigma + (x-\mu)/\sigma)].$$

*We shall write $\mathbf{X} \sim NL(\mu, \sigma^2, \alpha, \beta)$ to indicate that a random variable $\mathbf{X}$ has such a distribution.*

Reed (2004) showed that the NL distribution arises as the convolution of a normal distribution and an asymmetric Laplace, i.e. $\mathbf{X} \sim \mathrm{NL}(\mu, \sigma^2, \alpha, \beta)$ can be represented as

$$\mathbf{X} \stackrel{d}{=} \mathbf{Z} + \mathbf{W} \qquad (1.2.1)$$

where $\mathbf{Z}$ and $\mathbf{W}$ are independent random variables with $\mathbf{Z} \sim \mathbf{N}(\mu, \sigma^2)$ and $\mathbf{W}$ following an asymmetric Laplace distribution with pdf

$$f_{\mathbf{W}}(w) = \begin{cases} \frac{\alpha\beta}{\alpha+\beta}\exp(\beta w) & \text{for } w \le 0 \\ \frac{\alpha\beta}{\alpha+\beta}\exp(-\alpha w) & \text{for } w > 0. \end{cases}$$

Kotz, Kozubowski and Podgórski (2001) proved that a Laplace random variable can be represented as the difference between two exponentially distributed variates. Therefore, it follows from the relation (1.2.1) that a NL($\mu$, $\sigma^2$, $\alpha$, $\beta$) random variable can be expressed as

$$\mathbf{X} \stackrel{d}{=} \mu + \sigma\mathbf{Z} + (1/\alpha)\mathbf{E}_1 - (1/\beta)\mathbf{E}_2 \tag{1.2.2}$$

where $\mathbf{E}_1$, $\mathbf{E}_2$ are independent standard exponential random variables and $\mathbf{Z}$ is a standard normal random variable independent of $\mathbf{E}_1$ and $\mathbf{E}_2$.

This provides a convenient way to simulate pseudo-random numbers from the NL distribution.

Reed (2004) established the following properties of the NL distribution:

• In comparison with the $\mathbf{N}(\mu, \sigma^2)$ distribution, the NL($\mu$, $\sigma^2$, $\alpha$) distribution will always have more weight in the tails, in the sense that for $x$ suitably small, $x \to -\infty$, $F(x) > \Phi((x-\mu)/\sigma)$, while for $x$ suitably large $1-F(x) > 1-\Phi((x-\mu)/\sigma)$.

The parameters $\alpha$ and $\beta$ determine the behavior in the left and right tails respectively. In the case $\alpha = \beta$ it is symmetric and bell-shaped, occupying an intermediate position between a normal and a Laplace distribution.

• The NL distribution is closed under linear transformation.

Precisely, if $\mathbf{X} \sim$ NL($\mu$, $\sigma^2$, $\alpha$, $\beta$) and $a$ and $b$ are any constants, then

$$a\mathbf{X} + b \sim NL(a\mu + b, a^2\sigma^2, \alpha/a, \beta/a).$$

• The NL distribution is infinitely divisible.

• The NL distribution is not closed under the convolution operation, i.e. sums of independent NL random variables do not themselves follow NL distributions.

• It is the distribution of the stopped (or observed) state of a Brownian motion with normally distributed starting value if the stopping hazard rate is constant. This follows from the fact that the state of such Brownian motion with fixed (non-random) initial state after an exponentially distributed time follows an asymmetric Laplace distribution.

Thus for example, if the logarithmic price of a stock or other financial asset followed Brownian motion, as has been widely assumed, the log(price) at the time of the first trade on a fixed day $n$, say, could be expected to follow a distribution close to a normal-Laplace. This is because the log(price) at the start of day $n$ would be normally distributed, while under the assumption that trades on day $n$ occur in a Poisson process, the time until the first trade would be exponentially distributed.

The double Pareto-lognormal (dPlN) distribution is related to the normal-Laplace distribution in the same way as the lognormal is related to the normal, i.e. a random variable $\mathbf{X}$ for which $\log \mathbf{X} \sim \mathrm{NL}(\mu, \sigma^2, \alpha, \beta)$ is defined as following the double Pareto-lognormal distribution. As such it could be termed the "log normal-Laplace".

Reed and Jorgensen (2004) showed that the dPlN distribution arises as that of the state of a geometric Brownian motion (GBM), with lognormally distributed initial state, after an exponentially distributed length of time (or equivalently as the distribution of the stopped state of such a GBM with constant stopping rate). They discussed the application of the dPlN distribution in modelling the size distributions of various phenomena, including incomes and earnings, human settlement sizes, oil-field volumes, particle sizes and stock price returns.

Huberman and Adamic (1999) and Mitzenmacher (2001) suggested that the dPlN distibution can be a candidate for the size distribution of World-Wide Web sites and computer files.

The generalized normal Laplace distribution is introduced by the following definition.

**Definition 1.2.2.** *The generalized normal Laplace (GNL) distribution is defined as that of a random variable* $\mathbf{X}$ *with characteristic function:*

$$\phi(t) \stackrel{def}{=} E[\exp(it\mathbf{X})] = \left[\frac{\alpha\beta\exp(\mu it - \sigma^2 t^2/2)}{(\alpha - it)(\beta + it)}\right]^{\rho} \tag{1.2.3}$$

*where* $\alpha, \beta, \rho$ *and* $\sigma$ *are positive parameters* , $-\infty < \mu < +\infty$ *and* $t$ *is a real number. We shall write* $\mathbf{X} \sim GNL(\mu,\ \sigma^2,\ \alpha,\ \beta,\ \rho)$ *to indicate that the random variable* $\mathbf{X}$ *follows such a distribution.*

The characteristic function $\phi(t)$ given by (1.2.3) can be written as:

$$\begin{aligned} \phi(t) &= \exp(\rho\mu it - \rho\sigma^2 t^2/2)\left[\frac{\alpha}{\alpha - it}\right]^{\rho}\left[\frac{\beta}{\beta + it}\right]^{\rho} \\ &= \exp(\rho\mu it)\exp(-\rho\sigma^2 t^2/2)\,(1 - it/\alpha)^{-\rho}\,(1 + it/\beta)^{-\rho}. \end{aligned} \tag{1.2.4}$$

Consider the random variable $\mathbf{Z}_1$ defined as $\mathbf{Z}_1 = \sigma\sqrt{\rho}\ \mathbf{Y}$, where $\mathbf{Y}$ has a normal distribution with mean 0 and variance 1, that is $\mathbf{Y} \sim \mathbf{N}(0,\ 1)$ and $\sigma$, $\rho$ positive. Then the characteristic function of the random variable $\mathbf{Z}_1$ is

$$\begin{aligned} \phi_{\mathbf{Z}_1}(t) &\stackrel{def}{=} E[\exp(it\mathbf{Z}_1)] = E[\exp(it\sigma\sqrt{\rho}\ \mathbf{Y})] \\ &= \phi_{\mathbf{Y}}[t\sigma\sqrt{\rho}] = \exp(-t^2\sigma^2\rho/2) \end{aligned} \tag{1.2.5}$$

since $\mathbf{Y} \sim \mathbf{N}(0,\ 1)$ and its characteristic function, using relation (1.1.1), is

$$\phi_{\mathbf{Y}}(t) = \exp(-t^2/2).$$

Consider now the random variables $\mathbf{Z}_2$ and $\mathbf{Z}_3$ defined as $\mathbf{Z}_2 = (1/\alpha)\mathbf{G}_1$ and $\mathbf{Z}_3 = (-1/\beta)\mathbf{G}_2$, where $\alpha$, $\beta$ are positive and $\mathbf{G}_1$, $\mathbf{G}_2$ follow a gamma distribution with shape parameter $\rho > 0$ and scale parameter $\theta = 1$, that is $\mathbf{G}_1,\ \mathbf{G}_2 \sim$ **Gamma**$(\rho,\ 1)$. Using the relation (1.1.2), the characteristic functions of $\mathbf{G}_1$ and $\mathbf{G}_2$ are given by

$$\phi_{\mathbf{G}_1}(t) = \phi_{\mathbf{G}_2}(t) = (1 - it)^{-\rho}, \quad t \in \mathbf{R}.$$

Thus the characteristic functions of $\mathbf{Z}_2$ and $\mathbf{Z}_3$ are obtained as:

$$\phi_{\mathbf{Z}_2}(t) \stackrel{def}{=} E[\exp(it\mathbf{Z}_2)] = E[\exp(i\frac{t}{\alpha}\mathbf{G}_1)] = \phi_{\mathbf{G}_1}(t/\alpha) = (1 - it/\alpha)^{-\rho}, \tag{1.2.6}$$

$$\phi_{\mathbf{Z}_3}(t) \stackrel{def}{=} E[\exp(it\mathbf{Z}_3)] = E[\exp(-i\frac{t}{\beta}\mathbf{G}_2)] = \phi_{\mathbf{G}_2}(-t/\beta) = (1 + it/\beta)^{-\rho}, \tag{1.2.7}$$

with $t \in \mathbf{R}$.

Using theorem 1.1.4., theorem 1.1.6. and the relations (1.2.5), (1.2.6) and (1.2.7), we can conclude from the representation (1.2.4) of the characteristic function of a random variable $\mathbf{X}$ which follows the generalized normal Laplace (GNL) distribution that $\mathbf{X}$ can be represented as:

$$\mathbf{X} \stackrel{d}{=} \rho\mu + \sigma\sqrt{\rho}\,\mathbf{Y} + (1/\alpha)\mathbf{G}_1 - (1/\beta)\mathbf{G}_2 \tag{1.2.8}$$

where $\mathbf{Y}$, $\mathbf{G}_1$ and $\mathbf{G}_2$ are independent random variables with $\mathbf{Y} \sim \mathbf{N}(0,\,1)$, i.e. with probability density function $f(x) = \frac{1}{\sqrt{2\pi}}\exp(-x^2/2),\ x \in \mathbf{R}$ and $\mathbf{G}_1$, $\mathbf{G}_2$ gamma random variables with shape parameter $\rho$ and scale parameter 1, i.e. with probability density function $g(x) = \frac{1}{\Gamma(\rho)}x^{\rho-1}\exp(-x),\ x > 0$.

The representation (1.2.8) provides a straightforward way to generate pseudo-random deviates following a generalized normal Laplace (GNL) distribution.

Using the representation (1.2.2) of the NL distribution and (1.2.8) of the GNL distribution, we can conclude that the GNL is an extension of the NL distribution.

Properties of the GNL distribution such as the infinitely divisibility, closure under convolution operation and formulas for the moments of the GNL are investigated in the following.

**Definition 1.2.3.** *A random variable $\mathbf{X}$ has an infinitely divisible distribution if and only if, for each $n$, there exist independent, identically distributed random variables $\mathbf{X}_{n,k}$, $1 \le k \le n$, such that $\mathbf{X} \stackrel{d}{=} \sum_{k=1}^{n} \mathbf{X}_{n,k}$ for all $n$, or, equivalently, if and only if*

$$\phi_{\mathbf{X}}(t) = (\phi_{\mathbf{X}_{n,1}}(t))^n \ \textit{for all } n.$$

**Remark 1.2.1.** *An inspection of some of the most familiar characteristic functions tells us that, for example, the normal, the gamma, the Cauchy, the Poisson and the degenerate distributions all are infinitely divisible as it is shown by Feller (1971).*

**Proposition 1.2.1.** *If $\mathbf{X}$ and $\mathbf{Y}$ are independent, infinitely divisible random variables, then, for any $a, b \in \mathbf{R}$, so is $a\mathbf{X} + b\mathbf{Y}$.*

PROOF. Since $\mathbf{X}$ and $\mathbf{Y}$ are independent, infinitely divisible random variables, for every $n$ there exist independent, identically distributed random variables $\mathbf{X}_{n,k}$, $1 \leq k \leq n$, and $\mathbf{Y}_{n,k}$, $1 \leq k \leq n$,, such that $\mathbf{X} \stackrel{d}{=} \sum_{k=1}^{n} \mathbf{X}_{n,k}$ and $\mathbf{Y} \stackrel{d}{=} \sum_{k=1}^{n} \mathbf{Y}_{n,k}$, so that

$$a\mathbf{X} + b\mathbf{Y} \stackrel{d}{=} a\sum_{k=1}^{n} \mathbf{X}_{n,k} + b\sum_{k=1}^{n} \mathbf{Y}_{n,k} \stackrel{d}{=} \sum_{k=1}^{n} (a\mathbf{X}_{n,k} + b\mathbf{Y}_{n,k}),$$

or, alternatively, via characteristic functions (using theorem 1.1.4.),

$$\phi_{a\mathbf{X}+b\mathbf{Y}}(t) = \phi_{\mathbf{X}}(at)\phi_{\mathbf{Y}}(bt) = \left(\phi_{\mathbf{X}_{n,1}}(at)\right)^n \left(\phi_{\mathbf{Y}_{n,1}}(bt)\right)^n$$

$$= \left(\phi_{\mathbf{X}_{n,1}}(at)\phi_{\mathbf{Y}_{n,1}}(bt)\right)^n = \left(\phi_{a\mathbf{X}_{n,1}+b\mathbf{Y}_{n,1}}(t)\right)^n.$$

Thus, $a\mathbf{X} + b\mathbf{Y}$ is infinitely divisible. □

Using now the remark 1.2.1., proposition 1.2.1. and the representation (1.2.8) of the GNL random variable, we can conclude that the GNL distribution is infinitely divisible.

The next proposition establishes that the $n$-fold convolution of a GNL random variable also follows a GNL distribution.

**Proposition 1.2.2.** *The sum of n independent and identically distributed GNL($\mu$, $\sigma^2$, $\alpha$, $\beta$, $\rho$) random variable has a GNL($\mu$, $\sigma^2$, $\alpha$, $\beta$, $n\rho$) distribution.*

PROOF. Let $\mathbf{X}_1, \mathbf{X}_2, ...., \mathbf{X}_n$ be independent and identically distributed GNL($\mu$, $\sigma^2$, $\alpha$, $\beta$, $\rho$) random variables and consider $\mathbf{S}_n = \mathbf{X}_1 + \mathbf{X}_2 + .... + \mathbf{X}_n$. Then, by using theorem 1.1.4. and relation (1.2.3), we have

$$\phi_{\mathbf{S}_n}(t) = \phi_{\mathbf{X}_1}(t)\phi_{\mathbf{X}_2}(t)....\phi_{\mathbf{X}_n}(t) = (\phi_{\mathbf{X}_1}(t))^n = \left[\frac{\alpha\beta\exp(\mu i t - \sigma^2 t^2/2)}{(\alpha - it)(\beta + it)}\right]^{n\rho}$$

Therefore, $\mathbf{S}_n$ follows a GNL($\mu$, $\sigma^2$, $\alpha$, $\beta$, $n\rho$) distribution, by theorem 1.1.6. and relation (1.2.3). □

In the following, the mean, variance and the sample cumulants of the GNL distribution are calculated.

**Definition 1.2.4.** *The nth cumulant of a random variable* $\mathbf{X}$*, denoted* $k_n$*, is defined as the coefficient of* $t^n/n!$ *in the Taylor's series expansion (about* $t = 0$*) of the cumulant generating function of* $\mathbf{X}$*, that is* $\log[M_{\mathbf{X}}(t)]$*, where* $M_{\mathbf{X}}(t)$ *is the moment generating function of* $\mathbf{X}$ *defined as*

$$M_{\mathbf{X}}(t) = E[\exp(t\mathbf{X})],$$

*provided the expectation is finite for* $|t| < h$*, for some* $h > 0$.

**Proposition 1.2.3.** *Let* $\mathbf{X}$ *be a random variable.*
*a) If* $E|\mathbf{X}| < \infty$*, then* $k_1 = E[\mathbf{X}]$.
*b) If* $E[\mathbf{X}^2] < \infty$*, then* $k_2 = Var[\mathbf{X}]$.

PROOF. *a)*

$$E[\mathbf{X}] = \frac{d}{dt}M_{\mathbf{X}}(t)|_{t=0} = M'_{\mathbf{X}}(0)$$

and

$$k_1 = \frac{d}{dt}\log[M_{\mathbf{X}}(t)]|_{t=0} = \frac{M'_{\mathbf{X}}(t)}{M_{\mathbf{X}}(t)}|_{t=0} = M'_{\mathbf{X}}(0), \text{ as } M_{\mathbf{X}}(0) = 1$$

So, $k_1 = E[\mathbf{X}]$.
*b)*

$$E[\mathbf{X}^2] = \frac{d^2}{dt^2}M_{\mathbf{X}}(t)|_{t=0} = M''_{\mathbf{X}}(0)$$

and therefore,

$$Var[\mathbf{X}] = E[\mathbf{X}^2] - E[\mathbf{X}]^2 = M''_{\mathbf{X}}(0) - M'_{\mathbf{X}}(0)^2$$

Now, using $M_{\mathbf{X}}(0) = 1$ we obtain

$$k_2 = \frac{d^2}{dt^2}\log[M_{\mathbf{X}}(t)|_{t=0} = \frac{M''_{\mathbf{X}}(t)M_{\mathbf{X}}(t) - M'_{\mathbf{X}}(t)^2}{M_{\mathbf{X}}(t)}|_{t=0} = M''_{\mathbf{X}}(0) - M'_{\mathbf{X}}(0)^2$$

So, $k_2 = Var[\mathbf{X}]$. □

Using theorem 1.1.7. and theorem 1.1.5. we can derive from the representation (1.2.8) the mean, variance and the higher order cumulants $k_n$ for $n > 2$

of a random variable with the generalized normal Laplace (GNL) distribution. Relation (1.2.8) is:

$$\mathbf{X} \stackrel{d}{=} \rho\mu + \sigma\sqrt{\rho}\,\mathbf{Y} + (1/\alpha)\mathbf{G}_1 - (1/\beta)\mathbf{G}_2,$$

where $\mathbf{Y} \sim \mathbf{N}(0,\,1)$, $\mathbf{G}_1$, $\mathbf{G}_2 \sim \mathbf{Gamma}(\rho,\,1)$ and $\mathbf{Y}$, $\mathbf{G}_1$, $\mathbf{G}_2$ are independent random variables.

Therefore, we have:

$$E[\mathbf{X}] = \rho\mu + \sigma\sqrt{\rho} \times 0 + (1/\alpha)\rho \times 1 - (1/\beta)\rho \times 1$$

$$= \rho(\mu + 1/\alpha - 1/\beta),$$

since for a random variable $\mathbf{G} \sim \mathbf{Gamma}(\rho,\,\theta)$, its mean is equal to $\rho/\theta$.

Thus, according to proposition 1.2.3., we have

$$k_1 = E[\mathbf{X}] = \rho(\mu + 1/\alpha - 1/\beta).$$

For the second moment, we have

$$Var[\mathbf{X}] = \sigma^2\rho \times 1 + (1/\alpha^2)\rho \times 1^2 + (1/\beta^2)\rho \times 1^2$$

$$= \rho(\sigma^2 + 1/\alpha^2 + 1/\beta^2),$$

since for a random variable $\mathbf{G} \sim \mathbf{Gamma}(\rho,\,\theta)$, its variance is equal to $\rho/\theta^2$ and $\mathbf{Y}$, $\mathbf{G}_1$, $\mathbf{G}_2$ are independent random variables.

Thus, according to proposition 1.2.3., we have

$$k_2 = Var[\mathbf{X}] = \rho(\sigma^2 + 1/\alpha^2 + 1/\beta^2).$$

Also, by theorem 1.1.5. and examples 1 and 2 from section 1.1., we have:

$$M_{\mathbf{X}}(t) \stackrel{def}{=} E[\exp(t\mathbf{X})] = \exp(\rho\mu t) M_{\mathbf{Y}}(t\sigma\sqrt{\rho}) M_{\mathbf{G}_1}(t/\alpha) M_{\mathbf{G}_2}(-t/\beta)$$

$$= \exp(\rho\mu t)\exp(t^2\sigma^2\rho/2)\,(1 - t/\alpha)^{-\rho}\,(1 + t/\beta)^{-\rho}\,,\ -\beta < t < \alpha.$$

Therefore,

$$\log[M_{\mathbf{X}}(t)] = \rho\mu t + t^2\sigma^2\rho/2 - \rho\log(1 - t/\alpha) - \rho\log(1 + t/\beta)\,, -\beta < t < \alpha.$$

From the Taylor's series expansion of

$$-\rho\log(1 - t/\alpha) = \sum_{n=1}^{\infty} \rho(n-1)!t^n/\alpha^n n!$$

and

$$-\rho \log (1 + t/\beta) = \sum_{n=1}^{\infty} (-1)^n (n-1)! t^n / \beta^n n!$$

we obtain the higher order cumulants of $\mathbf{X}$ for $n > 2$:

$$k_n = \rho(n-1)! \left(\alpha^{-n} + (-1)^n \beta^{-n}\right), \; n > 2.$$

From the mean, variance and the cumulants $k_n, n > 2$ of $\mathbf{X} \sim \mathrm{GNL}(\mu,\, \sigma^2,\, \alpha,\, \beta,\, \rho)$ we conclude that the parameters $\mu$ and $\sigma^2$ influence the central location and spread of the distribution, while $\alpha$ and $\beta$ affect the lengths of the tails. The parameter $\rho$ affects all moments. In the following we calculate the coefficients of skewness:

$$\gamma_1 = \frac{k_3}{k_2^{3/2}} = \frac{2\rho\left(\alpha^{-3} - \beta^{-3}\right)}{\rho^{3/2}\left(\alpha^{-2} + \beta^{-2}\right)^{3/2}} = \frac{2\left(\alpha^{-3} - \beta^{-3}\right)}{\rho^{1/2}\left(\alpha^{-2} + \beta^{-2}\right)^{3/2}}$$

and kurtosis:

$$\gamma_2 = \frac{k_4 + 3k_2^2}{k_2^2} = \frac{6\rho\left(\alpha^{-4} + \beta^{-4}\right)}{\rho^2\left(\alpha^{-2} + \beta^{-2}\right)^2} + 3 = \frac{6\left(\alpha^{-4} + \beta^{-4}\right)}{\rho\left(\alpha^{-2} + \beta^{-2}\right)^2} + 3.$$

We note that both $\gamma_1$ and $\gamma_2 - 3$ decrease with increasing $\rho$ (and converge to zero as $\rho \to \infty$) with the shape of the distribution becoming more normal with increasing $\rho$ (exemplifying the central limit effect since the sum of $n$ independent and identically distributed $\mathrm{GNL}(\mu,\, \sigma^2,\, \alpha,\, \beta,\, \rho)$ random variables has a $\mathrm{GNL}(\mu,\, \sigma^2,\, \alpha,\, \beta,\, n\rho)$ distribution). When $\alpha = \beta$, the distribution is symmetric.
In the limiting case, $\alpha = \beta = \infty$ the relation (1.2.3) becomes

$$\phi(t) = exp(\rho\mu it - \rho\sigma^2 t^2/2)$$

and the GNL distribution reduces to a normal distribution.

A closed-form expression for the density has not been obtained except in the special case $\rho = 1$. In this case, based on the representations (1.2.2) and (1.2.8), the GNL distribution becomes the normal Laplace (NL) distribution.

The Black-Scholes theory of option pricing was originally based on the assumption that asset prices follow geometric Brownian motion (GBM). For such a process the logarithmic returns $(\log(P_{t+1}/P_t))$ on the price $P_t$ are independent identically distributed (iid) normal random variables. However, it has been recognized for some time now that the logarithmic returns do not behave quite like this, particularly over short intervals. Empirical distributions of the logarithmic returns in high-frequency data usually exhibit excess kurtosis with more probability mass near the origin and in the tails and less in the flanks than would occur for normally distributed data. Furthermore Rydberg (2000) showed that the degree of excess kurtosis increases as the sampling interval decreases. In addition skewness can sometimes be present. To accomodate for these facts new models for price movement based on Lévy motion have been developed. For example, Schoutens (2003) presented such models.

For any infinitely divisible distribution a Lévy process can be constructed whose increments follow the given distribution.

While the NL distribution is infinitely divisible, it is not closed under the convolution operation, i.e. sums of independent NL random variables do not themselves follow NL distributions. We saw that this closure property holds for the GNL and the advantage of this is that for such a class of distributions one can construct a Brownian-Laplace motion for which the increments follow the given distribution.

Reed (2005) defines Brownian-Laplace motion as a Lévy process which has both continuous (Brownian) and discontinuous (Laplace motion) components. The increments of the process follow a generalized normal Laplace (GNL) distribution which exhibits positive kurtosis and can be either symmetrical or exhibit skewness. The degree of kurtosis in the increments increases as the time between observations decreases.

Based on these properties, Reed (2005) proposed Brownian-Laplace motion as a good candidate model for the movement of the logarithmic price of a financial asset. He derived an option pricing formula for European call options and used this formula to calculate numerically the value of such an option both using nominal parameter values (to explore its independence upon them) and those obtained as estimates from real stock price data.

## 1.3. Estimation

We saw earlier that in the case $\rho = 1$, the generalized normal Laplace (GNL) distribution becomes the normal Laplace (NL) distribution.

For the NL distribution, maximum likelihood estimation of parameters can be carried out numerically since there is a closed-form expression for the probability density function (pdf). In fact it is shown in Reed and Jorgensen (2003) how one can estimate $\mu$ analytically and then maximize numerically the concentrated (profile) log-likelihood over the remaining three parameters.

Another approach, also discussed by Reed and Jorgensen (2003), uses the EM-algorithm (considering an NL random variable as the sum of normal and Laplace components, with one regarded as missing data).

Things are more difficult for the GNL distribution, since the lack of a closed-form expression for the generalized normal Laplace density means a similar lack for the likelihood function. This presents difficulties for estimation by maximum likelihood (ML). However it may be possible to obtain ML estimates of the parameters of the distribution using the EM-algorithm and the representation (1.2.8), but to date this has not been accomplished.

An alternative method of estimation is the method of moments. While method of moments estimates are consistent and asymptotically normal, they are not generally efficient (not achieving the Cramér-Rao bound), even asymptotically.

A further problem is the difficulty in restricting the parameter space (e.g. for the generalized normal Laplace distribution, the parameters $\alpha$, $\beta$, $\rho$ and $\sigma^2$ must be positive), since the moment equations may not lead to solutions in the restricted space.

In the case of a symmetric generalized normal Laplace distribution ($\alpha = \beta$) the method of moments estimators of the four parameters can be found analytically as follows.

In section 1.2., we established the formulas for the cumulants of the GNL distribution. These formulas are

$$k_1 = \rho(\mu + \alpha^{-1} - \beta^{-1})$$

$$k_2 = \rho(\sigma^2 + \alpha^{-2} + \beta^{-2})$$

and

$$k_n = \rho(n-1)!(\alpha^{-n} + (-1)^n \beta^{-n}), \quad for\ n \geq 3.$$

We note in this case, when $\alpha = \beta$, that for $n$ odd ($n \geq 3$), the cumulants $k_n$ are equal to zero. Therefore, in order to calculate the estimates of the four parameters $\widehat{\alpha} = \widehat{\beta}$, $\widehat{\rho}$, $\widehat{\mu}$ and $\widehat{\sigma^2}$, we need to solve the equations obtained by setting the sample cumulants $k_1$, $k_2$, $k_4$ and $k_6$ equal to their theoretical counterparts.

These equations are

$$k_1 = \rho\mu \tag{1.3.1}$$

$$k_2 = \rho(\sigma^2 + 2\alpha^{-2}) \tag{1.3.2}$$

$$k_4 = 12\rho\alpha^{-4} \tag{1.3.3}$$

$$k_6 = 240\rho\alpha^{-6} \tag{1.3.4}$$

Using equations (1.3.3) and (1.3.4) and the fact that $\alpha > 0$, we obtain

$$\alpha = \sqrt{20\frac{k_4}{k_6}}.$$

Now, using the same equations but in the following form:

$$k_4^3 = 12^3\rho^3\alpha^{-12}$$

and

$$k_6^2 = 240^2\rho^2\alpha^{-12},$$

we obtain

$$\rho = \frac{100}{3}\frac{k_4^3}{k_6^2}.$$

From equations (1.3.1) and (1.3.2), using the results for $\alpha$ and $\rho$ we obtain the values of $\mu$ and $\sigma^2$ as

$$\mu = \frac{k_1}{\rho}$$

and

$$\sigma^2 = \frac{k_2}{\rho} - \frac{2}{\alpha^2}$$

Thus, the estimates of the four parameters in the case of a symmetric GNL are

$$\widehat{\alpha} = \widehat{\beta} = \sqrt{20\frac{k_4}{k_6}}$$

$$\widehat{\rho} = \frac{100}{3} \frac{k_4^3}{k_6^2}$$

$$\widehat{\mu} = \frac{k_1}{\widehat{\rho}}$$

and

$$\widehat{\sigma}^2 = \frac{k_2}{\widehat{\rho}} - \frac{2}{\widehat{\alpha}^2},$$

where $k_i$ ($i$=1, 2, 4, 6) is the $i$th sample cumulant obtained either from the sample moments about zero, using formulas established by Kendall and Stuart (1969) or from the Taylor's series expansion of the sample cumulant generating function $\log(\frac{1}{n}\sum_{i=1}^{n}\exp(tx_i))$.

For the asymmetric generalized normal Laplace distribution with five parameters, numerical methods must be used to solve the moment equations produced by setting the first five sample cumulants equal to their theoretical counterparts, using the established formulas in section 1.2. that is,

$$k_1 = \rho(\mu + \alpha^{-1} - \beta^{-1}) \tag{1.3.5}$$

$$k_2 = \rho(\sigma^2 + \alpha^{-2} + \beta^{-2}) \tag{1.3.6}$$

$$k_3 = 2\rho(\alpha^{-3} - \beta^{-3}) \tag{1.3.7}$$

$$k_4 = 6\rho(\alpha^{-4} + \beta^{-4}) \tag{1.3.8}$$

$$k_5 = 24\rho(\alpha^{-5} - \beta^{-5}) \tag{1.3.9}$$

which can be reduced to a pair of nonlinear equations in two variables $\alpha$ and $\beta$ e.g.

$$12k_3(\alpha^{-5} - \beta^{-5}) = k_5(\alpha^{-3} - \beta^{-3})$$

$$3k_3(\alpha^{-4} + \beta^{-4}) = k_4(\alpha^{-3} - \beta^{-3})$$

and then the solutions for the other three parameters are obtained by substitution.

One drawback with the method of moments is that it is difficult to impose constraints on parameters (such as requiring estimates of $\alpha$, $\beta$, $\rho$, and $\sigma^2$ be positive) and estimates which are unsatisfactory in this respect may sometimes occur.

# Chapter 2

## QUADRATIC TRANSFORM DISTANCE ESTIMATORS

Luong and Thompson (1987) introduced estimators based on a "quadratic transform distance" and derived the asymptotic properties of these estimators.

The main justification for studying quadratic distance estimators is that they provide an alternative to maximum-likelihood estimation in cases when the likelihood function is difficult to compute. Examples include the fitting of stable distributions via the empirical characteristic function, and the estimation of parameters of mixture models like the "Poisson sum of normals" and "normal with gamma variance" models considered by Epps and Pulley (1985).

For the positive stable laws, Luong and Doray (2006) developed inference methods based on a special quadratic distance using negative moments.

Discrete distributions for modelling count data have been used in many fields of research such as actuarial science, biometry and economics, but many new distributions have complicated probability mass functions (pmf).

Luong and Doray (2002) developed general quadratic distance methods for discrete distributions definable recursively. They showed that these methods are applicable to many families of discrete distributions including those with complicated pmfs. They also concluded that the quadratic distance estimator protects against a certain form of misspecification of the distribution, which makes the maximum likelihood estimator biased, while keeping the quadratic distance estimator unbiased.

In situations where maximum-likelihood estimation is problematic, as where the likelihood function is unbounded or multi-modal, such quadratic transform methods as the method of moments, the minimum-chi-squared method, and variants of these can be practically and theoretically more tractable.

In this chapter, estimators based on a "quadratic transform distance" are introduced. Their influence functions are derived, and sufficient conditions for their consistency and asymptotic normality are given. Corresponding goodness-of-fit tests in cases of simple and composite null hypotheses are developed.

## 2.1. Definition

Let $x_1, x_2, ...., x_n$ be independently and identically distributed univariate observations from a cumulative distribution function $F_\theta$, where $\theta = [\theta_1, \theta_2, ...., \theta_m] \in \Theta \subseteq \mathbf{R}^\mathbf{m}$ and let $F_n$ be the sample cumulative distribution function.

Let $\overline{\Theta}$ denote the closure of $\Theta$ in $\mathbf{R}^\mathbf{m}$ and $'$ denote the transpose of a vector.

The class "quadratic transform distance" (QTD) is the class of "distances" $d$ of the form

$$d(F_n, F_\theta) = [z(F_n) - z(F_\theta)]' Q(F_\theta)[z(F_n) - z(F_\theta)],$$

where

(1) for a fixed set of functions $h_j, j = 1, 2, ....k$, the transform of a distribution $F$ is a vector

$$z(F) = [z_1(F), z_2(F), ...., z_k(F)]',$$

where

$$z_j(F) = \int_{-\infty}^{\infty} h_j(x) dF(x), \qquad j = 1, 2, ...., k,$$

if the $h_j$ are integrable with respect to $F$;

(2) for any $\theta \in \overline{\Theta}$ the matrix $Q(F_\theta)$ is a defined, continuous, symmetric, and positive definite matrix of dimension $k \times k$;

(3) under each $F_\theta$ the variance-covariance matrix of $[h_1(x), h_2(x), ...., h_k(x)]$ exists; we denote it by $\Sigma(F_\theta) = (\sigma_{ij})_{1\leq i,j\leq k}$, where

$$\sigma_{ij} = \int_{-\infty}^{\infty} h_i(x)h_j(x)dF_\theta(x) - z_i(F_\theta)z_j(F_\theta)$$

and note that it is also the variance-covariance matrix of $\sqrt{n}z'(F_n)$ under $F_\theta$.

For simplicity, in what follows we shall use $z(\theta)$, $z_n$, $Q(\theta)$, $\Sigma(\theta)$ for $z(F_\theta)$, $z(F_n)$, $Q(F_\theta)$, $\Sigma(F_\theta)$ respectively.

Thus the class QTD consists of distances of the form

$$d(F_n, F_\theta) = [z_n - z(\theta)]'Q(\theta)[z_n - z(\theta)]. \tag{2.1.1}$$

Corresponding to the distance $d$, the QTD estimator is the value $\widehat{\theta}_n$ of $\theta$ which minimizes (2.1.1) over $\Theta$, if this minimum is attained.

## 2.2. Consistency and Asymptotic Normality of the Quadratic Distance Estimator

A sufficient condition for consistency is established in the following proposition.

**Proposition 2.2.1.** *With the notations introduced in section 2.1., let $Q(\theta)$ be continuous in $\theta$, and symmetric and positive definite for all $\theta \in \overline{\Theta}$ (the closure of the parameter space in $\mathbf{R}^{\mathbf{m}}$). Let $\widehat{\theta}_n$ be an estimator minimizing a distance of form (2.1.1).*

*If $k > m$, and the function $\theta \longrightarrow z(\theta)$ from $\Theta$ to $\mathbf{R}^{\mathbf{k}}$ has a continuous inverse at $\theta_0$ (in the sense that whenever $z(\theta_n) \longrightarrow z(\theta_0)$ we have $\theta_n \longrightarrow \theta_0$), then $\widehat{\theta}_n \xrightarrow{P} \theta_0$, where $F_{\theta_0}$ is the true cumulative distribution function of the $x_i$ and $\xrightarrow{P}$ denotes convergence in probability.*

Proof. By definition of $\widehat{\theta}_n$, we have

$$0 \leq [z_n - z(\widehat{\theta}_n)]'Q(\widehat{\theta}_n)[z_n - z(\widehat{\theta}_n)] \leq [z_n - z(\theta_0)]'Q(\theta_0)[z_n - z(\theta_0)]. \tag{2.2.1}$$

Since $z_n \xrightarrow{\text{P}} z(\theta_0)$ by the weak law of large numbers, the right hand side of (2.2.1) tends to 0 in probability, and hence so does the middle term. This can only happen if $z_n - z(\widehat{\theta}_n) \xrightarrow{\text{P}} 0$; but this implies $z(\widehat{\theta}_n) \xrightarrow{\text{P}} z(\theta_0)$, which implies $\widehat{\theta}_n \xrightarrow{\text{P}} \theta_0$ by the continuous-inverse assumption. This completes the proof. $\square$

The following proposition will help to establish the asymptotic normality of the quadratic distance estimator.

**Proposition 2.2.2.** *Suppose that $k > m$, that $\widehat{\theta}_n$ is consistent in the sense of the preceding discussion, and that the following conditions hold, where* P *refers to probability under $F(\theta_0)$:*

*(i) $Q$ is symmetric, continuous, and positive definite (and has partial derivatives with respect to its $\theta_j$- components);*

*(ii) with probability converging to 1 the estimate $\widehat{\theta}_n$ satisfies the m-dimensional system*

$$\sqrt{n}\frac{\partial}{\partial\theta}\left[[z_n - z(\widehat{\theta}_n)]'Q(\widehat{\theta}_n)[z_n - z(\widehat{\theta}_n)]\right] = 0; \tag{2.2.2}$$

*(iii) each $\partial z_i(\theta)/\partial\theta_j$ is continuous at $\theta_0$, $i = 1, 2, ...., k$ and $j = 1, 2, ...., m$;*

*(iv) the $k \times m$ matrix $S(\theta)$ whose $i, j$th element is $\partial z_i(\theta)/\partial\theta_j$ has rank $m$ for $\theta = \theta_0$;*

*(v) $w_n' = [\partial Q(\theta)/\partial\theta_1, \partial Q(\theta)/\partial\theta_2, ...., \partial Q(\theta)/\partial\theta_m]_{\theta=\widehat{\theta}_n}$ is bounded in probability.*

*Then*

$$\sqrt{n}(\widehat{\theta}_n - \theta_0) = (S'QS)^{-1}S'Q\sqrt{n}[z_n - z(\theta_0)] + o_p(1), \tag{2.2.3}$$

*where $S = S(\theta_0)$, $Q = Q(\theta_0)$ and $o_p(1)$ denotes an expression converging to 0 in probability.*

PROOF. With probability $1 - o_p(1)$, we have $\widehat{\theta}_n$ satisfying (2.2.2), which is equivalent to the following $m$-dimensional system:

$$\sqrt{n}(\frac{\partial}{\partial\theta}z'(\widehat{\theta}_n))Q(\widehat{\theta}_n)[z_n - z(\widehat{\theta}_n)] - \frac{\sqrt{n}}{2}[z_n - z(\widehat{\theta}_n)]'\frac{\partial Q}{\partial\theta}(\widehat{\theta}_n)[z_n - z(\widehat{\theta}_n)] = 0 \quad (2.2.4)$$

By condition *(v)* above and the inequality (2.2.1), the system (2.2.4) reduces to

$$\sqrt{n}(\frac{\partial}{\partial\theta}z'(\widehat{\theta}_n))Q(\widehat{\theta}_n)[z_n - z(\widehat{\theta}_n)] + o_p(1) = 0.$$

More simply, we can express this system as

$$\sqrt{n}S'(\widehat{\theta}_n)Q(\widehat{\theta}_n)[z_n - z(\widehat{\theta}_n)] + o_p(1) = 0. \quad (2.2.5)$$

By the mean-value theorem, we have

$$z(\widehat{\theta}_n) = z(\theta_0) + [S(\theta_0) + o_p(1)](\widehat{\theta}_n - \theta_0),$$

which implies

$$\sqrt{n}S'(\widehat{\theta}_n)Q(\widehat{\theta}_n)[z_n - z(\widehat{\theta}_n)]$$
$$= \sqrt{n}S'(\widehat{\theta}_n)Q(\widehat{\theta}_n)[z_n - z(\theta_0)] - S'(\widehat{\theta}_n)Q(\widehat{\theta}_n)[S(\theta_0) + o_p(1)]\sqrt{n}(\widehat{\theta}_n - \theta_0),$$

so that we obtain

$$S'(\widehat{\theta}_n)Q(\widehat{\theta}_n)[S(\theta_0) + o_p(1)]\sqrt{n}(\widehat{\theta}_n - \theta_0) = \sqrt{n}S'(\widehat{\theta}_n)Q(\widehat{\theta}_n)[z_n - z(\theta_0)] + o_p(1)$$

by using (2.2.5).

Now using the fact that $\widehat{\theta}_n$ is consistent, that is, $\widehat{\theta}_n \xrightarrow{P} \theta_0$ it results

$$\sqrt{n}(\widehat{\theta}_n - \theta_0) = (S'QS)^{-1}S'Q\sqrt{n}\ [z_n - z(\theta_0)] + o_p(1),$$

where $S = S(\theta_0)$ and $Q = Q(\theta_0)$. □

**Remark 2.2.1.** *Luong and Thompson (1987) showed that $\sqrt{n}\ [z_n - z(\theta_0)] \xrightarrow{\mathcal{L}} N(0, \Sigma)$, where $\Sigma$ is the variance-covariance matrix of $\sqrt{n}\ [z_n - z(\theta_0)]$ under the hypothesis $\theta = \theta_0$. Therefore, it follows from the above proposition that*

$$\sqrt{n}(\widehat{\theta}_n - \theta_0) \xrightarrow{\mathcal{L}} N(0, \Sigma_1),$$

*where $\Sigma_1 = (S'QS)^{-1}S'Q\Sigma QS(S'QS)^{-1}$ and $\xrightarrow{\mathcal{L}}$ denotes convergence in law.*

*Clearly $\Sigma_1$ will take a simpler form if $S'Q\Sigma QS = S'QS$.*

*Also, if $\Sigma$ is invertible, it is easily seen that an optimal choice of $Q$ in the sense of minimizing the norm of the variance-covariance matrix $\Sigma_1$ is $Q_0 = \Sigma^{-1}$, and then*

$$\Sigma_1 = (S'\Sigma^{-1}S)^{-1}S'\Sigma^{-1}\Sigma\Sigma^{-1}S(S'\Sigma^{-1}S)^{-1}$$

$$= (S'\Sigma^{-1}S)^{-1}(S'\Sigma^{-1}S)(S'\Sigma^{-1}S)^{-1} = (S'\Sigma^{-1}S)^{-1}$$

*Thus $\Sigma_1 = (S'\Sigma^{-1}S)^{-1}$.*

*For the norm, we will use the determinant of the variance-covariance matrix.*

## 2.3. Influence functions for the quadratic distance estimators

Luong and Thompson (1987) derived the influence functions for estimators in the class "quadratic transform distance" (QTD) for the special case where $\theta$ is one-dimensional, the general case being $\theta$ $m$-dimensional.

In order to show this result, the functional formulation is presented and based on this formulation, the influence functions are defined and derived for the quadratic distance estimators. It will be concluded that the quadratic distance estimators are robust when the functions $h_1(x)$, $h_2(x)$,....,$h_k(x)$ are bounded functions of $x$.

Under the conditions that $d(F_n, F_\theta)$ attains its minimum at an interior point of $\Theta$ and that $z(\theta)$ and $Q(\theta)$ are differentiable, the estimator $\widehat{\theta}_n$ minimizing the distance (2.1.1), namely

$$d(F_n, F_\theta) = [z_n - z(\theta)]'Q(\theta)[z_n - z(\theta)],$$

may also be defined as a root of the $m$-dimensional system of estimating equations:

$$\frac{\partial}{\partial\theta}(d(F_n, F_\theta)) = 0.$$

Specifically, this system becomes:

$$\frac{\partial}{\partial\theta}([z_n - z(\theta)]'Q(\theta)[z_n - z(\theta)]) = 0.$$

This enables us to express $\widehat{\theta}_n$ in its "functional formulation" as $\widehat{\theta}_n = T(F_n)$, where $T(G)$ is that value of $\theta$ for which $d(G, F_\theta)$ is minimized. The estimate $\widehat{\theta}_n$ is then Fisher consistent in the sense that $T$ satisfies the relation $T(F_\theta) = \theta$. Implicitly, $T(G)$ is a root of

$$\frac{\partial}{\partial\theta}\left([z(G) - z(\theta)]'Q(\theta)[z(G) - z(\theta)]\right) = 0. \tag{2.3.1}$$

Consider now

$$F_\lambda = (1-\lambda)F + \lambda\delta_x,$$

where $\delta_x$ is the degenerate distribution at $x$ and $0 < \lambda < 1$.

The influence function of Hampel (1968, 1974) is defined as

$$IC_{T,F}(x) = \frac{\partial T(F_\lambda)}{\partial\lambda}|_{\lambda=0}\ ,$$

where $T$ is the functional defined above.

This function measures the influence on $T(F)$ of contaminating $F$ by mixing it with the singular distribution at points $x$, and hence can be used to measure the response of $T(F)$ to outlier contamination.

If $H(\theta,\lambda)$ is the left-hand side of (2.3.1) when $G = F_\lambda$, we have $H(T(F_\lambda),\lambda) = 0$. Thus

$$\frac{\partial H}{\partial\theta}(\theta,\lambda)|_{\theta=T(F),\lambda=0} \times \frac{\partial T(F_\lambda)}{\partial\lambda}|_{\lambda=0} + \frac{\partial H}{\partial\lambda}(\theta,\lambda)|_{\theta=T(F),\lambda=0} = 0,$$

so that

$$IC_{T,F}(x) = -\left(\frac{\partial H}{\partial\lambda}(\theta,\lambda)|_{\theta=T(F),\lambda=0}\right) / \left(\frac{\partial H}{\partial\theta}(\theta,\lambda)|_{\theta=T(F),\lambda=0}\right). \tag{2.3.2}$$

The numerator is:

$$\frac{\partial H}{\partial\lambda}(\theta,\lambda)|_{\theta=T(F),\lambda=0}$$

$$= \left(-2\frac{\partial z'(\theta)}{\partial\theta}Q(\theta)\frac{\partial z(F_\lambda)}{\partial\lambda} + 2[z(F_\lambda) - z(\theta)]'\frac{\partial Q(\theta)}{\partial\theta}\frac{\partial z(F_\lambda)}{\partial\lambda}\right)|_{\theta=T(F),\lambda=0} \tag{2.3.3}$$

In the case of $F = F_\theta$, when the strict parametric model holds, only the first term is present.

The denominator is expressible similarly:

$$\frac{\partial H}{\partial\theta}(\theta,\lambda)|_{\theta=T(F),\lambda=0}$$

$$= (2\frac{\partial z'(\theta)}{\partial\theta}Q(\theta)\frac{\partial z(\theta)}{\partial\theta} - 2\frac{\partial^2 z'(\theta)}{\partial\theta^2}Q(\theta)[z(F_\lambda) - z(\theta)]$$

$$-2\frac{\partial z'(\theta)}{\partial\theta}\frac{\partial Q(\theta)}{\partial\theta}[z(F_\lambda) - z(\theta)]$$

$$+[z(F_\lambda) - z(\theta)]'\frac{\partial^2 Q(\theta)}{\partial\theta^2}[z(F_\lambda) - z(\theta)])|_{\theta=T(F),\lambda=0}. \qquad (2.3.4)$$

Since

$$z'(F_\lambda) = \left[\int_{-\infty}^{\infty} h_1(x)dF_\lambda(x), \int_{-\infty}^{\infty} h_2(x)dF_\lambda(x), ...., \int_{-\infty}^{\infty} h_k(x)dF_\lambda(x)\right],$$

$$z'(F) = \left[\int_{-\infty}^{\infty} h_1(x)dF(x), \int_{-\infty}^{\infty} h_2(x)dF(x), ...., \int_{-\infty}^{\infty} h_k(x)dF(x)\right]$$

and $F_\lambda = (1-\lambda)F + \lambda\delta_x$, we have

$$z'(F_\lambda) = (1-\lambda)z'(F) + z'(\delta_x),$$

where $z'(\delta_x) = [h_1(x), h_2(x), ...., h_k(x)]$.

Thus,

$$\frac{\partial z'(F_\lambda)}{\partial\lambda}|_{\lambda=0} = z'(\delta_x) - z'(F). \qquad (2.3.5)$$

It follows that the influence function for any QTD estimator will have the form:

$$IC_{T,F}(x) = \frac{2\left(\frac{\partial z'(\theta)}{\partial\theta}Q(\theta) - [z(F) - z(\theta)]'\frac{\partial Q(\theta)}{\partial\theta}\right)|_{\theta=T(F)}\ [z(\delta_x) - z(F)]}{\frac{\partial H}{\partial\theta}(\theta,\lambda)|_{\theta=T(F),\lambda=0}} \qquad (2.3.6)$$

according to the relations (2.3.2), (2.3.3) and (2.3.5).

If $Q$ is not a function of $\theta$ that is, $Q(\theta) = Q$, according also to (2.3.4) the formula (2.3.6) becomes:

$$IC_{T,F}(x) = \frac{2\frac{\partial z'(\theta)}{\partial\theta}|_{\theta=T(F)}\ Q[z(\delta_x) - z(F)]}{\left(2\frac{\partial z'(\theta)}{\partial\theta}Q\frac{\partial z(\theta)}{\partial\theta} - 2\frac{\partial^2 z'(\theta)}{\partial\theta^2}Q[z(F_\lambda) - z(\theta)]\right)|_{\theta=T(F)}}$$

$$= \frac{\frac{\partial z'(\theta)}{\partial\theta}|_{\theta=T(F)}\ Q[z(\delta_x) - z(F)]}{\left(\frac{\partial z'(\theta)}{\partial\theta}Q\frac{\partial z(\theta)}{\partial\theta} - \frac{\partial^2 z'(\theta)}{\partial\theta^2}Q[z(F_\lambda) - z(\theta)]\right)|_{\theta=T(F)}}.$$

For $Q$ depending on $\theta$, but for $F = F_\theta$, we have $z(F) = z(\theta)$ and therefore,

$$IC_{T,F_\theta}(x) = \frac{2\frac{\partial z'(\theta)}{\partial\theta}Q(\theta)[z(\delta_x) - z(\theta)]}{2\frac{\partial z'(\theta)}{\partial\theta}Q(\theta)\frac{\partial z(\theta)}{\partial\theta}}$$

$$= \frac{\frac{\partial z'(\theta)}{\partial \theta} Q(\theta)[z(\delta_x) - z(\theta)]}{\frac{\partial z'(\theta)}{\partial \theta} Q(\theta) \frac{\partial z(\theta)}{\partial \theta}}.$$

**Remark 2.3.1.** *We note that for fixed $F$ the influence function $IC_{T,F}(x)$ is equivalent to a linear combination of the functions $h_1(x), h_2(x), ...., h_k(x)$.*

*Thus if the $h_j(x)$, $j = 1, 2, ...., k$ are bounded functions of $x$, then clearly $IC_{T,F}(x)$ is bounded for fixed $F$, and the corresponding quadratic distance estimator is robust in this sense (see Hampel (1974)).*

## 2.4. Goodness of fit tests

Luong and Thompson (1987) developed corresponding goodness-of-fit tests in cases of simple and composite null hypotheses and limiting chi-square distributions are obtained for distance-based test statistics. In the case of a composite hypothesis the limiting distribution of the estimated quadratic distance is chi-square (up to a constant) only when the quadratic distance estimator is used to estimate the parameter.

Therefore, aiming at constructing test statistics with a limiting chi-square distribution, some mathematical results are needed and these are presented in the subsection 2.4.1.

In subsections 2.4.2 and 2.4.3, the goodness of fit tests for the simple and composite hypotheses respectively are presented.

### 2.4.1. Some mathematical results

The following theorem also used in Luong and Thompson (1987) is needed; its proof can be found in Moore (1977, 1978) or Rao (1973).

**Theorem 2.4.1.** *Suppose that the random vector $\mathbf{Y}$ of dimension $p$ is $N_p(0, \Sigma)$ and $C$ is any $p \times p$ symmetric positive semi-definite matrix, then the quadratic form $\mathbf{Y}'C\mathbf{Y}$ is chi-square distributed with $r$ degrees of freedom if $\Sigma C$ is idempotent and $trace(\Sigma C) = r$.*

*(The same result holds asymptotically if $C$ is replaced by a consistent estimate $\widehat{C}$*

*and* $\mathbf{Y} \xrightarrow{\mathcal{L}} N_p(0, \Sigma)$).

The following remark used in Luong and Thompson (1987) is needed.

**Remark 2.4.1.** *In order to have* $\Sigma C$ *idempotent, it suffices to have either*

$$(i)\ \Sigma C \Sigma = \Sigma, \ \text{i.e. } C = \Sigma^-, \ C \text{ being a generalized inverse of } \Sigma,$$

*or*

$$(ii)\ \Sigma C \Sigma = C, \ \text{i.e. } \Sigma = C^-, \ \Sigma \text{ being a generalized inverse of } C.$$

### 2.4.2. Goodness of fit for the simple hypothesis

To test the null hypothesis $\mathcal{H}_0$: $F = F_{\theta_0}$ against all alternatives, the following test statistic based on a distance in a quadratic distance class can be used:

$$nd(F_n, F_{\theta_0}) = n[z_n - z]' Q [z_n - z] = v_n' Q v_n,$$

where $z = z(\theta_0)$ and $v_n = \sqrt{n}\,[z_n - z]$.

The goal is to choose $Q$ so that $v_n' Q v_n$ will have a chi-square limiting distribution under the null hypothesis.

As it was mentioned in remark 2.2.1., $v_n \xrightarrow{\mathcal{L}} N(0, \Sigma)$, where $\Sigma$ is the variance-covariance matrix of the $\sqrt{n}\,[z_n - z(\theta_0)]$ under the hypothesis $\theta = \theta_0$, by remark 2.4.1 and theorem 2.4.1, it suffices to choose $Q$ to be symmetric and a generalized inverse of $\Sigma$.

The number of degrees of freedom of the limiting chi-square distribution is given by

$$r = r(\Sigma Q) = trace(\Sigma Q).$$

Thus, if $\Sigma$ is invertible and $Q = \Sigma^{-1}$, then $v_n' Q v_n \xrightarrow{\mathcal{L}} \chi^2(k)$ since

$$trace(\Sigma Q) = trace(\Sigma \Sigma^{-1}) = trace(I_k) = k.$$

### 2.4.3. Goodness of fit for the composite hypothesis

In order to construct a test for the composite hypothesis $\mathcal{H}_0$: $F \in \{F_\theta,\ \theta \in \Theta\}$ ($\Theta$ is the full parameter space), we must estimate $\theta$. The following theorem indicates how to choose an estimator for $\theta$ so that the test statistic will have a limiting chi-square distribution.

**Theorem 2.4.2.** *Let $\widehat{\theta}_n$ be a minimum distance estimator based on the quadratic transform distance*

$$d(F_n, F_\theta) = [z_n - z(\theta)]' Q(\theta) [z_n - z(\theta)].$$

*Suppose the assumptions of proposition 2.2.2. hold, and let $Q$ be such that the matrix $\Sigma_2 Q$ is idempotent, where $\Sigma_2$ is given by (2.4.3) below.*
*Then if $v_n(\widehat{\theta}_n) = \sqrt{n}\ [z_n - z(\widehat{\theta}_n)]$, the test statistic*

$$D^2(\widehat{\theta}_n) = nd(F_n, F_{\widehat{\theta}_n}) = v_n'(\widehat{\theta}_n) Q(\widehat{\theta}_n) v_n(\widehat{\theta}_n) \tag{2.4.1}$$

*will have a limiting chi-squared distribution under $\mathcal{H}_0$, with $r = trace(\Sigma_2 Q)$ degrees of freedom. Note that $\Sigma_2 Q$ is idempotent if for example $Q = \Sigma^{-1}$, where $\Sigma$ is the variance-covariance matrix of $\sqrt{n}\ [z_n - z(\theta_0)]$ under $F_{\theta_0}$.*

PROOF. Using a Taylor's series expansion, we have

$$v_n(\widehat{\theta}_n) = v_n - S\sqrt{n}(\widehat{\theta}_n - \theta_0) + o_p(1),$$

where $S$ is the matrix with $(i, j)$th entry $\partial z_i / \partial \theta_j$ evaluated at $\theta_0$ and $o_p(1)$ denotes an expression converging to 0 in probability. By proposition 2.2.2., we have

$$\begin{aligned}\sqrt{n}(\widehat{\theta}_n - \theta_0) &= (S'QS)^{-1} S'Q\sqrt{n}\ [z_n - z(\theta_0)] + o_p(1) \\ &= (S'QS)^{-1} S'Q v_n + o_p(1),\end{aligned}$$

which gives us

$$\begin{aligned}v_n(\widehat{\theta}_n) &= v_n - S(S'QS)^{-1} S'Q v_n + o_p(1) \\ &= \left[I_k - S(S'QS)^{-1} S'Q\right] v_n + o_p(1).\end{aligned} \tag{2.4.2}$$

Since $v_n \xrightarrow{\mathcal{L}} N(0, \Sigma)$, we have $v_n(\widehat{\theta}_n) \xrightarrow{\mathcal{L}} N(0, \Sigma_2)$, where

$$\Sigma_2 = \left[I_k - S(S'QS)^{-1}S'Q\right] \Sigma \left[I_k - QS(S'QS)^{-1}S'\right]. \qquad (2.4.3)$$

Using theorem 2.4.1., $v_n'(\widehat{\theta}_n)Q(\widehat{\theta}_n)v_n(\widehat{\theta}_n)$ will have a limiting chi-square distribution if $\Sigma_2 Q$ is idempotent with $r = r(\Sigma_2 Q) = trace(\Sigma_2 Q)$ degrees of freedom.

If $Q = \Sigma^{-1}$, we obtain

$$\begin{aligned}
\Sigma_2 Q &= \left[I_k - S(S'QS)^{-1}S'Q\right] \Sigma \left[I_k - QS(S'QS)^{-1}S'\right] Q \\
&= \left[I_k - S(S'QS)^{-1}S'Q\right] \left[I_k - S(S'QS)^{-1}S'Q\right] \\
&= I_k - S(S'QS)^{-1}S'Q - S(S'QS)^{-1}S'Q + S(S'QS)^{-1}S'QS(S'QS)^{-1}S'Q \\
&= I_k - S(S'QS)^{-1}S'Q.
\end{aligned}$$

Thus the idempotency of $\Sigma_2 Q$ follows from the idempotency of the second term since

$$\begin{aligned}
\left(S(S'QS)^{-1}S'Q\right)^2 &= S(S'QS)^{-1}(S'QS)(S'QS)^{-1}S'Q \\
&= S(S'QS)^{-1}S'Q.
\end{aligned}$$

By Rao (1973), the rank of $\Sigma_2 Q$ is given by

$$rank(\Sigma_2 Q) = k - rank(S\left(S'QS)^{-1}S'Q\right) = k - m$$

Therefore, in the case $Q = \Sigma^{-1}$, the test statistic

$$D^2(\widehat{\theta}_n) = nd(F_n, F_{\widehat{\theta}_n}) = v_n'(\widehat{\theta}_n)Q(\widehat{\theta}_n)v_n(\widehat{\theta}_n)$$

will have a limiting chi-square distribution under $\mathcal{H}_0$, with $r = k - m$ degrees of freedom. $\square$

# Chapter 3

---

# INFERENCE FOR THE PARAMETERS OF THE GNL DISTRIBUTION USING QUADRATIC DISTANCE METHOD

In this chapter, the quadratic distance estimator based on the characteristic function is investigated in order to estimate the parameters of the generalized normal Laplace distribution. This is a special distance within the class of quadratic transform distances presented in section 2.1.. First, we present the empirical characteristic function and some properties.
We establish the expression for the variance-covariance matrix of the errors between the empirical and characteristic functions.

Properties of the quadratic distance estimator (QDE) such as robustness, consistency and asymptotic normality are given in this chapter. For model testing, the test statistic based on the quadratic distance follows a unique chi-square distribution under the composite null hypothesis, which makes the test statistics easy to use.

Quadratic distance methods can be viewed as alternative methods to classical ones such as the method of moments, the maximum likelihood and the minimum chi-square methods.

## 3.1. Empirical characteristic function and its properties

Let $\mathbf{X}, \mathbf{X}_1, \mathbf{X}_2, \mathbf{X}_3, .....$be independent and identically distributed random variables with distribution function $F(x) = P(\mathbf{X} \leq x)$ and characteristic function

$$\phi(t) = E[\exp(it\mathbf{X})] = \int_{-\infty}^{\infty} \exp(itx)dF(x), \ -\infty < t < \infty.$$

Many proposed statistical procedures for sequences of independent and identically distributed (i.i.d.) random variables may be thought of as based on the empirical cumulative distribution function $F_n(x) = N(x)/n$, where $N(x)$ is the number of $\mathbf{X}_j \leq x$ with $1 \leq j \leq n$, for example, procedures based on statistics of a Kolmogorov-Smirnov type.

In view of the one-to-one correspondence between distribution functions and characteristic functions, it seems natural to investigate procedures based on the empirical characteristic function (sample characteristic function) defined as

$$\phi_n(t) = \int_{-\infty}^{\infty} \exp(itx)dF_n(x) = \frac{1}{n}\sum_{k=1}^{n} \exp(it\mathbf{X}_k), \quad -\infty < t < \infty.$$

This function is the characteristic function of the empirical distribution of the data, which assigns probability $1/n$ to each observation and is seen to be an average of $n$ independent processes of the type $\exp(it\mathbf{X})$.

The range of problems to which the empirical characteristic function seems applicable appears to be quite wide. This is because the Fourier-Stieltjes transformation often results in easy translation of properties that are important in problems of inference.

**Proposition 3.1.1.** *Let $\mathbf{X}_1, \mathbf{X}_2, ...., \mathbf{X}_n$ be independent and identically distributed random variables with the characteristic function $\phi(t)$. Then,*

$$E[\phi_n(t)] = \phi(t)$$

*and*

$$\phi_n(t) \to \phi(t) \textit{ almost surely as } n \to \infty.$$

PROOF.

$$E[\phi_n(t)] = E[\frac{1}{n}\sum_{k=1}^{n}\exp(it\mathbf{X}_k)] = \frac{1}{n}\sum_{k=1}^{n}E[\exp(it\mathbf{X}_k)] = \frac{1}{n}\sum_{k=1}^{n}\phi(t) = \phi(t),$$

as $\mathbf{X}_1, \mathbf{X}_2, ...., \mathbf{X}_n$ are identically distributed.

Since $\mathbf{X}_1, \mathbf{X}_2, ...., \mathbf{X}_n$ are independent and identically distributed, we know from theorem 1.1.3. that the same is true for $\exp(it\mathbf{X}_1), \exp(it\mathbf{X}_2), ...., \exp(it\mathbf{X}_n)$.

Therefore, it follows by the strong law of large numbers that $\phi_n(t)$ converges almost surely to $\phi(t) = E[\exp(it\mathbf{X}_1)]$. □

Feuerverger and Mureika (1977) proved the following result:

If $\phi_n(t)$ is the empirical characteristic function corresponding to a characteristic function $\phi(t)$ then we have, for fixed $T < \infty$, the convergence

$$sup_{|t|\leq T}|\phi_n(t) - \phi(t)| \to 0 \textit{ almost surely as } n \to \infty$$

(a consequence of the Glivenko-Cantelli and the P. Lévy theorems).

They explained that this uniform convergence cannot generally take place on the whole line. However, they showed that it does hold on the whole line under some general restrictions such as $F$ is purely discrete.

Such results naturally lead one to expect that the empirical characteristic function has valuable statistical applications. Consequently, estimators based on the sample characteristic function are usually strongly consistent.

The following proposition establishes formulas in order to construct the variance-covariance matrix of the vector $[\mathrm{Re}\phi_n(t_1), ...., \mathrm{Re}\phi_n(t_k), \mathrm{Im}\phi_n(t_1), ...., \mathrm{Im}\phi_n(t_k)]$ for a finite collection of points $t_1, t_2, \ldots, t_k$. We prove these formulas using the properties of the empirical characteristic function.

**Proposition 3.1.2.** *Let* $\mathbf{X}_1, \mathbf{X}_2, ...., \mathbf{X}_n$ *be independent and identically distributed random variables with the characteristic function* $\phi(t)$. *Then,*

$$E(\mathrm{Re}\phi_n(t), \mathrm{Im}\phi_n(t)) = (\mathrm{Re}\phi(t), \mathrm{Im}\phi(t)) \tag{3.1.1}$$

$$Cov(\mathrm{Re}\phi_n(t), \mathrm{Re}\phi_n(s)) = \frac{1}{2n}[\mathrm{Re}\phi(t+s) + \mathrm{Re}\phi(t-s) - 2\mathrm{Re}\phi(t)\mathrm{Re}\phi(s)] \quad (3.1.2)$$

$$Cov(\mathrm{Im}\phi_n(t), \mathrm{Im}\phi_n(s)) = \frac{1}{2n}[\mathrm{Re}\phi(t-s) - \mathrm{Re}\phi(t+s) - 2\mathrm{Im}\phi(t)\mathrm{Im}\phi(s)] \quad (3.1.3)$$

$$Cov(\mathrm{Re}\phi_n(t), \mathrm{Im}\phi_n(s)) = \frac{1}{2n}[\mathrm{Im}\phi(t+s) - \mathrm{Im}\phi(t-s) - 2\mathrm{Re}\phi(t)\mathrm{Im}\phi(s)]. \quad (3.1.4)$$

PROOF. Since $\phi_n(t) = \mathrm{Re}\phi_n(t) + i\mathrm{Im}\phi_n(t)$ and $E[\phi_n(t)] = E[\mathrm{Re}\phi_n(t)] + iE[\mathrm{Im}\phi_n(t)]$, it results immediately that

$$E(\mathrm{Re}\phi_n(t), \mathrm{Im}\phi_n(t)) = (\mathrm{Re}\phi(t), \mathrm{Im}\phi(t))$$

by proposition 3.1.1..

In order to obtain the relations (3.1.2), (3.1.3) and (3.1.4) first we obtain an expression for $E[\phi_n(t)\phi_n(s)]$:

$$E[\phi_n(t)\phi_n(s)] = E[\frac{1}{n^2}\left(\exp(i(t+s)\mathbf{X}_1) + \exp(i(t+s)\mathbf{X}_2) + .... + \exp(i(t+s)\mathbf{X}_n)\right)$$

$$+\frac{1}{n^2}\exp(it\mathbf{X}_1)\left(\exp(is\mathbf{X}_2) + .... + \exp(is\mathbf{X}_n)\right) + ....$$

$$+\frac{1}{n^2}\exp(it\mathbf{X}_n)\left(\exp(is\mathbf{X}_1) + .... + \exp(is\mathbf{X}_{n-1})\right)].$$

Since $\mathbf{X}_1, \mathbf{X}_2, ...., \mathbf{X}_n$ are independent and identically distributed random variables with the characteristic function $\phi(t)$, we have

$$E[\phi_n(t)\phi_n(s)] = \frac{1}{n^2}n\phi(t+s) + \frac{1}{n^2}n(n-1)\phi(t)\phi(s)$$

$$E[\phi_n(t)\phi_n(s)] = \frac{1}{n}\phi(t+s) + \frac{n-1}{n}\phi(t)\phi(s). \qquad (i)$$

Since $\phi_n(-s) = \frac{1}{n}\sum_{k=1}^{n}\exp(i(-s)\mathbf{X}_k) = \overline{\phi_n(s)}$, we have also

$$E[\phi_n(t)\overline{\phi_n(s)}] = \frac{1}{n}\phi(t-s) + \frac{n-1}{n}\phi(t)\overline{\phi(s)} \qquad (ii)$$

Writing $\phi_n(t) = \mathrm{Re}\phi_n(t) + i\mathrm{Im}\phi_n(t)$, $\phi_n(s) = \mathrm{Re}\phi_n(s) + i\mathrm{Im}\phi_n(s)$ and using relations $(i)$ and $(ii)$ we obtain by identifying the real and imaginary parts the following relations:

- $E[\mathrm{Re}\phi_n(t)\mathrm{Re}\phi_n(s)] - E[\mathrm{Im}\phi_n(t)\mathrm{Im}\phi_n(s)]$

$$= \frac{1}{n}\mathrm{Re}\phi(t+s) + \frac{n-1}{n}\mathrm{Re}\phi(t)\mathrm{Re}\phi(s) - \frac{n-1}{n}\mathrm{Im}\phi(t)\mathrm{Im}\phi(s)$$

- $E[\text{Re}\phi_n(t)\text{Re}\phi_n(s)] + E[\text{Im}\phi_n(t)\text{Im}\phi_n(s)]$
$$= \frac{1}{n}\text{Re}\phi(t-s) + \frac{n-1}{n}\text{Re}\phi(t)\text{Re}\phi(s) + \frac{n-1}{n}\text{Im}\phi(t)\text{Im}\phi(s)$$
- $E[\text{Re}\phi_n(t)\text{Im}\phi_n(s)] + E[\text{Im}\phi_n(t)\text{Re}\phi_n(s)]$
$$= \frac{1}{n}\text{Im}\phi(t+s) + \frac{n-1}{n}\text{Re}\phi(t)\text{Im}\phi(s) + \frac{n-1}{n}\text{Im}\phi(t)\text{Re}\phi(s)$$
- $-E[\text{Re}\phi_n(t)\text{Im}\phi_n(s)] + E[\text{Im}\phi_n(t)\text{Re}\phi_n(s)]$
$$= \frac{1}{n}\text{Im}\phi(t-s) - \frac{n-1}{n}\text{Re}\phi(t)\text{Im}\phi(s) + \frac{n-1}{n}\text{Im}\phi(t)\text{Re}\phi(s).$$

From the first 2 relations we obtain

$$E[\text{Re}\phi_n(t)\text{Re}\phi_n(s)] = \frac{1}{2n}\text{Re}\phi(t+s) + \frac{1}{2n}\text{Re}\phi(t-s) + \frac{n-1}{n}\text{Re}\phi(t)\text{Re}\phi(s) \quad (iii)$$

$$E[\text{Im}\phi_n(t)\text{Im}\phi_n(s)] = \frac{1}{2n}\text{Re}\phi(t-s) - \frac{1}{2n}\text{Re}\phi(t+s) + \frac{n-1}{n}\text{Im}\phi(t)\text{Im}\phi(s), \quad (iv)$$

and from the last 2 relations we obtain

$$E[\text{Re}\phi_n(t)\text{Im}\phi_n(s)] = \frac{1}{2n}\text{Im}\phi(t+s) - \frac{1}{2n}\text{Im}\phi(t-s) + \frac{n-1}{n}\text{Re}\phi(t)\text{Im}\phi(s) \quad (v)$$

Using now relations $(iii)$, $(iv)$, $(v)$ and (3.1.1) the desired relations are obtained as follows:

$$Cov(\text{Re}\phi_n(t), \text{Re}\phi_n(s)) = E[\text{Re}\phi_n(t)\text{Re}\phi_n(s)] - E[\text{Re}\phi_n(t)][E[\text{Re}\phi_n(s)]$$
$$= \frac{1}{2n}[\text{Re}\phi(t+s) + \text{Re}\phi(t-s) - 2\text{Re}\phi(t)\text{Re}\phi(s)]$$
$$Cov(\text{Im}\phi_n(t), \text{Im}\phi_n(s)) = E[\text{Im}\phi_n(t)\text{Im}\phi_n(s)] - E[\text{Im}\phi_n(t)][E[\text{Im}\phi_n(s)]$$
$$= \frac{1}{2n}[\text{Re}\phi(t-s) - \text{Re}\phi(t+s) - 2\text{Im}\phi(t)\text{Im}\phi(s)]$$
$$Cov(\text{Re}\phi_n(t), \text{Im}\phi_n(s)) = E[\text{Re}\phi_n(t)\text{Im}\phi_n(s)] - E[\text{Re}\phi_n(t)][E[\text{Im}\phi_n(s)]$$
$$= \frac{1}{2n}[\text{Im}\phi(t+s) - \text{Im}\phi(t-s) - 2\text{Re}\phi(t)\text{Im}\phi(s)].$$

□

In the following, it is shown that, for finite collections $t_1, t_2, ...., t_m$, the finite-dimensional distributions of processes $\mathbf{Y}_n(t_j) = \sqrt{n}\,(\phi_n(t_j) - \phi(t_j))$ converge to those of a complex Gaussian process $\mathbf{Y}(\mathbf{t_j})$, for $j = 1, 2, \ldots, m$.

**Proposition 3.1.3.** *Consider* $\mathbf{Y}_n(t) = \sqrt{n}\,(\phi_n(t) - \phi(t))$ *as a random complex process in* $t$. *Then,*

*a)* $E[\mathbf{Y}_n(t)] = 0$;

*b)* $E[\mathbf{Y}_n(t)\mathbf{Y}_n(s)] = \phi(t+s) - \phi(t)\phi(s)$;

*c) The covariance structure of* $\mathbf{Y}_n(t)$ *is given by:*

$$Cov\,(\mathrm{Re}\mathbf{Y}_n(t), \mathrm{Re}\mathbf{Y}_n(s)) = \frac{1}{2}\left[\mathrm{Re}\phi(t+s) + \mathrm{Re}\phi(t-s)\right] - \mathrm{Re}\phi(t)\mathrm{Re}\phi(s) \quad (3.1.5)$$

$$Cov\,(\mathrm{Im}\mathbf{Y}_n(t), \mathrm{Im}\mathbf{Y}_n(s))$$
$$= \frac{1}{2}\left[-\mathrm{Re}\phi(t+s) + \mathrm{Re}\phi(t-s)\right] - \mathrm{Im}\phi(t)\mathrm{Im}\phi(s) \quad (3.1.6)$$

$$Cov\,(\mathrm{Re}\mathbf{Y}_n(t), \mathrm{Im}\mathbf{Y}_n(s)) = \frac{1}{2}\left[\mathrm{Im}\phi(t+s) - \mathrm{Im}\phi(t-s)\right] - \mathrm{Re}\phi(t)\mathrm{Im}\phi(s). \quad (3.1.7)$$

PROOF. *a)*

$$E[\mathbf{Y}_n(t)] = \sqrt{n}\,(E[\phi_n(t)] - \phi(t)) = 0$$

since $E[\phi_n(t)] = \phi(t)$ by proposition 3.1.1.

*b)*

$$E[\mathbf{Y}_n(t)\mathbf{Y}_n(s)] = n\,(E[\phi_n(t)\phi_n(s)] - \phi(s)E[\phi_n(t)] - \phi(t)E[\phi_n(s)] + \phi(t)\phi(s))$$

$$E[\mathbf{Y}_n(t)\mathbf{Y}_n(s)] = n\,(E[\phi_n(t)\phi_n(s)] - \phi(s)\phi(t) - \phi(t)\phi(s) + \phi(t)\phi(s))$$

$$E[\mathbf{Y}_n(t)\mathbf{Y}_n(s)] = n\,(E[\phi_n(t)\phi_n(s)] - \phi(t)\phi(s))$$

In the proof of proposition 3.1.2., the following relation was derived:

$$E[\phi_n(t)\phi_n(s)] = \frac{1}{n}\phi(t+s) + \frac{n-1}{n}\phi(t)\phi(s) \qquad (i)$$

Using this relation $(i)$ we have further

$$E[\mathbf{Y}_n(t)\mathbf{Y}_n(s)] = n\left(\frac{1}{n}\phi(t+s) + \frac{n-1}{n}\phi(t)\phi(s) - \phi(t)\phi(s)\right)$$
$$= \phi(t+s) - \phi(t)\phi(s).$$

*c)* Since $\mathbf{Y}_n(t) = \sqrt{n}\,(\phi_n(t) - \phi(t))$ we have

$$\mathrm{Re}\mathbf{Y}_n(t) = \sqrt{n}\,(\mathrm{Re}\phi_n(t) - \mathrm{Re}\phi(t))$$

and

$$\mathrm{Im}\mathbf{Y}_n(t) = \sqrt{n}\,(\mathrm{Im}\phi_n(t) - \mathrm{Im}\phi(t))\,.$$

So, by using relations (3.1.2), (3.1.3), (3.1.4) from proposition 3.1.2, the covariance structure of $\mathbf{Y}_n(t)$ is calculated as follows:

$$
\begin{aligned}
Cov\left(\mathrm{Re}\mathbf{Y}_n(t), \mathrm{Re}\mathbf{Y}_n(s)\right) &= Cov\left(\sqrt{n}(\mathrm{Re}\phi_n(t) - \mathrm{Re}\phi(t), \sqrt{n}(\mathrm{Re}\phi_n(s) - \mathrm{Re}\phi(s)\right) \\
&= nCov\left(\mathrm{Re}\phi_n(t), \mathrm{Re}\phi_n(s)\right) - nCov\left(\mathrm{Re}\phi_n(t), \mathrm{Re}\phi(s)\right) \\
&\quad -nCov\left(\mathrm{Re}\phi(t), \mathrm{Re}\phi_n(s)\right) + nCov\left(\mathrm{Re}\phi(t), \mathrm{Re}\phi(s)\right) \\
&= nCov\left(\mathrm{Re}\phi_n(t), \mathrm{Re}\phi_n(s)\right) \\
&= n\frac{1}{2n}\left[\mathrm{Re}\phi(t+s) + \mathrm{Re}\phi(t-s) - 2\mathrm{Re}\phi(t)\mathrm{Re}\phi(s)\right] \\
&= \frac{1}{2}\left[\mathrm{Re}\phi(t+s) + \mathrm{Re}\phi(t-s)\right] - \mathrm{Re}\phi(t)\mathrm{Re}\phi(s).
\end{aligned}
$$

$$
\begin{aligned}
Cov\left(\mathrm{Im}\mathbf{Y}_n(t), \mathrm{Im}\mathbf{Y}_n(s)\right) &= Cov\left(\sqrt{n}(\mathrm{Im}\phi_n(t) - \mathrm{Im}\phi(t)), \sqrt{n}(\mathrm{Im}\phi_n(s) - \mathrm{Im}\phi(s)\right) \\
&= nCov\left(\mathrm{Im}\phi_n(t), \mathrm{Im}\phi_n(s)\right) - nCov\left(\mathrm{Im}\phi_n(t), \mathrm{Im}\phi(s)\right) \\
&\quad -nCov\left(\mathrm{Im}\phi(t), \mathrm{Im}\phi_n(s)\right) + nCov\left(\mathrm{Im}\phi(t), \mathrm{Im}\phi(s)\right) \\
&= nCov\left(\mathrm{Im}\phi_n(t), \mathrm{Im}\phi_n(s)\right) \\
&= n\frac{1}{2n}\left[\mathrm{Re}\phi(t-s) - \mathrm{Re}\phi(t+s) - 2\mathrm{Im}\phi(t)\mathrm{Im}\phi(s)\right] \\
&= \frac{1}{2}\left[-\mathrm{Re}\phi(t+s) + \mathrm{Re}\phi(t-s)\right] - \mathrm{Im}\phi(t)\mathrm{Im}\phi(s).
\end{aligned}
$$

$$
\begin{aligned}
Cov\left(\mathrm{Re}\mathbf{Y}_n(t), \mathrm{Im}\mathbf{Y}_n(s)\right) &= Cov\left(\sqrt{n}(\mathrm{Re}\phi_n(t) - \mathrm{Re}\phi(t)), \sqrt{n}(\mathrm{Im}\phi_n(s) - \mathrm{Im}\phi(s)\right) \\
&= nCov\left(\mathrm{Re}\phi_n(t), \mathrm{Im}\phi_n(s)\right) - nCov\left(\mathrm{Re}\phi_n(t), \mathrm{Im}\phi(s)\right) \\
&\quad -nCov\left(\mathrm{Re}\phi(t), \mathrm{Im}\phi_n(s)\right) + nCov\left(\mathrm{Re}\phi(t), \mathrm{Im}\phi(s)\right) \\
&= nCov\left(\mathrm{Re}\phi_n(t), \mathrm{Im}\phi_n(s)\right) \\
&= n\frac{1}{2n}\left[\mathrm{Im}\phi(t+s) - \mathrm{Im}\phi(t-s) - 2\mathrm{Re}\phi(t)\mathrm{Im}\phi(s)\right] \\
&= \frac{1}{2}\left[\mathrm{Im}\phi(t+s) - \mathrm{Im}\phi(t-s)\right] - \mathrm{Re}\phi(t)\mathrm{Im}\phi(s).
\end{aligned}
$$

□

**Remark 3.1.1.** *Because $\phi_n(t)$ is an average of bounded processes it follows also, by means of the multidimensional central limit theorem, for finite collections $t_1, t_2, ...., t_m$ that $\mathbf{Y}_n(t_1), \mathbf{Y}_n(t_2), ...., \mathbf{Y}_n(t_m)$ converge in distribution to $\mathbf{Y}(t_1)$, $\mathbf{Y}(t_2)$, ....,$\mathbf{Y}(t_m)$ where $\mathbf{Y}(t) = \overline{\mathbf{Y}(-t)}$ is a complex Gaussian process with covariance structure identical to $\mathbf{Y}_n(t)$, namely*

$$E[\mathbf{Y}(t)] = 0$$

*and*

$$E[\mathbf{Y}(t)\mathbf{Y}(s)] = \phi(t+s) - \phi(t)\phi(s).$$

*More generally, Feuerverger and Mureika (1977) proved that the process $\mathbf{Y}_n(t)$ converges weakly to $\mathbf{Y}(t)$ in every finite interval.*

**Remark 3.1.2.** *The characteristic function behaves simply under shifts, scale changes and summation of independent variables; it allows an easy characterization of independence and of symmetry. It is therefore not difficult to suggest empirical characteristic function procedures for areas of inference such as testing for goodness of fit and parameter estimation.*

*The empirical characteristic function retains all information present in the sample and lends itself conveniently to computation. Also, there are situations where a characterization of some property or of a class of distributions exists in terms of characteristic functions. One example is the problem of inference on the parameters of the stable laws. Here traditional methods have not led to a solution and an empirical characteristic function approach seems likely to be a useful procedure. See for example Paulson, Halcomb and Leitch (1975).*

## 3.2. Quadratic distance estimator based on the characteristic function of the GNL distribution

Let $\mathbf{X}$ be a random variable which follows GNL distribution. According to definition 1.2.2., the characteristic function of $\mathbf{X}$ is given by

$$\phi(t) \stackrel{def}{=} E[\exp(it\mathbf{X})] = \left[\frac{\alpha\beta\exp(\mu it - \sigma^2 t^2/2)}{(\alpha - it)(\beta + it)}\right]^{\rho},$$

where $\alpha, \beta, \rho$ and $\sigma$ are positive parameters , $-\infty < \mu < +\infty$ and $t$ is a real number.

Also, consider the empirical characteristic function presented in section 3.1. that is,

$$\phi_n(t_j) = \frac{1}{n}\sum_{l=1}^{n}\exp(it_j\mathbf{X}_l), \; j = 1, 2, ...., k,$$

where $\mathbf{X}_1, \mathbf{X}_2, ...., \mathbf{X}_n$ are independent, identically distributed observations from the GNL distribution, $k$ and the points $t_1, t_2, ...., t_k$ are fixed and do not vary with the sample $n$.

Let us define the vectors

$$\mathbf{Z}_n = [\mathrm{Re}\phi_n(t_1), ...., \mathrm{Re}\phi_n(t_k), \mathrm{Im}\phi_n(t_1), ...., \mathrm{Im}\phi_n(t_k)]'$$

$$\mathbf{Z}(\theta) = [\mathrm{Re}\phi(t_1), ...., \mathrm{Re}\phi(t_k), \mathrm{Im}\phi(t_1), ...., \mathrm{Im}\phi(t_k)]',$$

where $\theta$ is the parameter vector $[\mu,\ \sigma^2,\ \alpha,\ \beta,\ \rho]'$. Consider the notation

$$\theta = [\mu, \sigma^2, \alpha, \beta, \rho]' = [\theta_1, \theta_2, \theta_3, \theta_4, \theta_5]'.$$

According to the quadratic transform distance presented in section 2.1., the quadratic distance estimator based on the characteristic function, denoted by $\widehat{\theta}$, is defined as the value of $\theta$ which minimizes the distance

$$d(\theta) = [\mathbf{Z}_n - \mathbf{Z}(\theta)]'Q(\theta)[\mathbf{Z}_n - \mathbf{Z}(\theta)], \tag{3.2.1}$$

where $Q(\theta)$ is a positive definite matrix which depends on $\theta$.

**Remark 3.2.1.** *In the framework of the quadratic transform distance estimators, the vector $(z(F))'$ is the $2k$-dimensional vector*

$$\left[\int_{-\infty}^{\infty} \cos(t_1 x) dF(x), ...., \int_{-\infty}^{\infty} \cos(t_k x) dF(x), \int_{-\infty}^{\infty} \sin(t_1 x) dF(x), ...., \int_{-\infty}^{\infty} \sin(t_k x) dF(x)\right],$$

*and clearly, the functions $h_j(x)$, $j = 1, 2, ...., 2k$ are:*

$$h_j(x) = \cos(t_j x) \ for \ j = 1, 2, ...., k$$

*and*

$$h_j(x) = \sin(t_{j-k} x) \ for \ j = k+1, ...., 2k.$$

*If $\phi_n(t) = \int \exp(itx) dF_n(x)$ is the empirical characteristic function, $\Sigma(\theta)$ is the variance-covariance matrix of the vector*

$$\sqrt{2n}\left[\mathrm{Re}\phi_n(t_1), ...., \mathrm{Re}\phi_n(t_k), \mathrm{Im}\phi_n(t_1), ...., \mathrm{Im}\phi_n(t_k)\right].$$

*So, according to proposition 3.1.2., the $\Sigma(\theta)$ is the $2k \times 2k$ symmetric matrix with elements:*

$$\sigma_{i,j}(\theta) = \begin{cases} \mathrm{Re}\phi(t_i + t_j) + \mathrm{Re}\phi(t_i - t_j) - 2\mathrm{Re}\phi(t_i)\mathrm{Re}\phi(t_j) & for\ 1 \le i,\ j \le k \\ \mathrm{Re}\phi(t_i - t_j) - \mathrm{Re}\phi(t_i + t_j) - 2\mathrm{Im}\phi(t_i)\mathrm{Im}\phi(t_j) & for\ k+1 \le i,\ j \le 2k \\ \mathrm{Im}\phi(t_i + t_j) - \mathrm{Im}\phi(t_i - t_j) - 2\mathrm{Re}\phi(t_i)\mathrm{Im}\phi(t_j) & for\ 1 \le i \le k,\ k+1 \le j \le 2k, \end{cases}$$

*where we define $t_i = t_{i-k}$ for $k+1 \le i \le 2k$.*

## 3.3. Consistency and asymptotic normality

Let $\theta_0$ be the true value of the parameter $\theta$. Using propositions 2.2.1., 2.2.2., the remark 2.2.1 and the remark 3.2.1., we can conclude that:

i) $\widehat{\theta}_n \xrightarrow{\mathrm{P}} \theta_0$, where $\xrightarrow{\mathrm{P}}$ denotes convergence in probability, i.e. the QDE $\widehat{\theta}$ is a consistent estimator of the parameter $\theta$,

ii) $\sqrt{n}(\widehat{\theta}_n - \theta_0) \xrightarrow{\mathcal{L}} N(0, \Sigma_1)$ where $\xrightarrow{\mathcal{L}}$ denotes convergence in law, with

a) $\Sigma_1 = (S'QS)^{-1} S'Q\Sigma QS(S'QS)^{-1}$,

b) $Q = Q(\theta_0)$,

c) $S = (\frac{\partial \mathbf{Z}_i(\theta)}{\partial \theta_j})$, a matrix of dimension $2k \times 5$, with $i = 1, 2, ...., 2k$ and $j = 1, 2, 3, 4, 5$, and $\mathbf{Z}_i(\theta) = \begin{cases} \text{Re}\phi(t_i) & \text{for } i = 1, 2, \ldots, k \\ \text{Im}\phi(t_{i-k}) & \text{for } i = k+1, \ldots, 2k, \end{cases}$

d) $\Sigma = (\sigma_{i,j})$ is the variance-covariance matrix of $\mathbf{Y}_n = \sqrt{n}\ [\mathbf{Z}_n - \mathbf{Z}(\theta_0)]$ under the hypothesis $\theta = \theta_0$.

iii) $Q(\theta)$ can be replaced by an estimate $\widehat{Q}$ in (3.2.1) and if $\widehat{Q} \xrightarrow{\text{P}} Q(\theta_0) = Q$, we have an estimator asymptotically equivalent to the one obtained by minimizing (3.2.1),

iv) The most efficient choice of $Q(\theta)$, is $\Sigma^{-1}(\theta)$, and with this choice of $Q(\theta)$, the corresponding QDE, denoted by $\widehat{\theta}^*$, has the property

$$\sqrt{n}(\widehat{\theta}^* - \theta_0) \xrightarrow{\mathcal{L}} N(0, (S'\Sigma^{-1}S)^{-1}).$$

Based on this property of asymptotic normality, at level $1 - \alpha$, the confidence intervals are given by the formulas:

$$\widehat{\theta}^* \pm Z_{\alpha/2} \times \sqrt{Var(\widehat{\theta}^*)},$$

where $\Phi(Z_{\alpha/2}) = 1 - \alpha/2$ and $\Phi(z)$ is the cdf of a standard normal random variable.

Therefore, we have the following algorithm for calculating the estimator. The most efficient estimator is $\widehat{\theta}$, obtained by choosing $Q = \Sigma^{-1}$. The easiest one to obtain is $\widetilde{\theta}$, obtained by choosing $Q = I$, the identity matrix. Despite the fact that $\widetilde{\theta}$ is less efficient, it can be used to estimate $\Sigma^{-1}$, by letting $\widehat{\Sigma}^{-1} = \Sigma^{-1}(\widetilde{\theta})$. We then can use $\widehat{\Sigma}^{-1}$ to obtain the first iteration for $\widehat{\theta}$ and this procedure can be repeated with $\Sigma^{-1}$ re-estimated at each step; $\widehat{\theta}$ is defined as the convergent vector value of the procedure.

We will focus on $\widehat{\theta}^*$, the most efficient QDE in this class of quadratic distance and note that the choice of points $t_1, t_2, ..., t_k$ affects $(S'\Sigma^{-1}S)^{-1}$. For fixed $k$, we have the design question for choice of the $t_1, t_2, ...., t_k$.

Feuerverger and McDunnough (1981) showed that by using a sufficiently extensive grid $t_j$, $j = 1, 2, ...., k$, $(S'\Sigma^{-1}S)^{-1}$ can be made arbitrarily close to the Cramér-Rao bound, so that the quadratic distance method can attain arbitrarily high asymptotic efficiency.

Also, the behavior of the stochastic processes $\mathrm{Re}\phi_n(t)$ and $\mathrm{Im}\phi_n(t)$ provides us with some guidance in the choice of the points $t_j$, $j = 1, 2, ...., k$. In the following, this behavior will be studied.

We saw in section 1.2., that for the generalized normal Laplace (GNL) distribution, the characteristic function is defined as

$$\phi(t) = \exp(\rho\mu it - \rho\sigma^2 t^2/2)\left[\frac{\alpha}{\alpha - it}\right]^{\rho}\left[\frac{\beta}{\beta + it}\right]^{\rho},$$

with $-\infty < t < \infty$, $\alpha, \beta, \sigma > 0$ and $-\infty < \mu < \infty$.

The real and imaginary parts of $\phi(t)$ are calculated from its trigonometric form as follows:

$$\phi(t) = \exp(-\rho\sigma^2 t^2/2)\exp(\rho\mu it)\,(1 - it/\alpha)^{-\rho}\,(1 + it/\beta)^{-\rho}$$

$$= \exp(-\rho\sigma^2 t^2/2)\left[\cos(\rho\mu t) + i\sin(\rho\mu t)\right]\left[(1 + it/\alpha)/(1 + t^2/\alpha^2)\right]^{\rho}\left[(1 - it/\beta)/(1 + t^2/\beta^2)\right]^{\rho}$$

$$\begin{aligned} &= \exp(-\rho\sigma^2 t^2/2)\left[\cos(\rho\mu t) + i\sin(\rho\mu t)\right]\left(1 + t^2/\alpha^2\right)^{-\rho/2}\left(1 + t^2/\beta^2\right)^{-\rho/2} \\ &\quad\times\left[\cos(\rho\theta_1) + i\sin(\rho\theta_1)\right]\left[\cos(\rho\theta_2) + +i\sin(\rho\theta_2)\right] \\ &= \exp(-\rho\sigma^2 t^2/2)\left(1 + t^2/\alpha^2\right)^{-\rho/2}\left(1 + t^2/\beta^2\right)^{-\rho/2} \\ &\quad\times\left[\cos\rho(\mu t + \theta_1 + \theta_2) + i\sin\rho(\mu t + \theta_1 + \theta_2)\right], \end{aligned} \tag{3.3.1}$$

where $\theta_1 = \arctan(t/\alpha)$ and $\theta_2 = \arctan(-t/\beta)$.

So, from relation (3.3.1) we have

$$\begin{aligned} \mathrm{Re}\phi(t) &= \exp(-\rho\sigma^2 t^2/2)\left(1 + t^2/\alpha^2\right)^{-\rho/2} \\ &\quad\times\left(1 + t^2/\beta^2\right)^{-\rho/2}\cos\rho(\mu t + \theta_1 + \theta_2) \end{aligned} \tag{3.3.2}$$

and

$$\begin{aligned} \mathrm{Im}\phi(t) &= \exp(-\rho\sigma^2 t^2/2)\left(1 + t^2/\alpha^2\right)^{-\rho/2} \\ &\quad\times\left(1 + t^2/\beta^2\right)^{-\rho/2}\sin\rho(\mu t + \theta_1 + \theta_2) \end{aligned} \tag{3.3.3}$$

From relation (3.1.2), we obtain

$$\begin{aligned} Var(\mathrm{Re}\phi_n(t)) &= \frac{1}{2n}\left[\mathrm{Re}\phi(2t) + \mathrm{Re}\phi(0) - 2(\mathrm{Re}\phi(t))^2\right] \\ &= \frac{1}{2n}\left[\mathrm{Re}\phi(2t) + 1 - 2(\mathrm{Re}\phi(t))^2\right] \end{aligned} \tag{3.3.4}$$

since $\phi(0) = 1$.

From relation (3.1.3), we obtain

$$Var(\mathrm{Im}\phi_n(t)) = \frac{1}{2n}\left[\mathrm{Re}\phi(0) - \mathrm{Re}\phi(2t) - 2(\mathrm{Im}\phi(t))^2\right]$$
$$= \frac{1}{2n}\left[1 - \mathrm{Re}\phi(2t) - 2(\mathrm{Im}\phi(t))^2\right] \tag{3.3.5}$$

since $\phi(0) = 1$.

Now, using relations (3.1.1), (3.3.2), (3.3.3), (3.3.4) and (3.3.5), it follows that, as $|t| \to \infty$,

$$E(\mathrm{Re}\phi_n(t), \mathrm{Im}\phi_n(t)) = (\mathrm{Re}\phi(t), \mathrm{Im}\phi(t))$$
$$= (\exp(-\rho\sigma^2 t^2/2)(1+t^2/\alpha^2)^{-\rho/2}(1+t^2/\beta^2)^{-\rho/2}\cos\rho(\mu t + \theta_1 + \theta_2),$$
$$\exp(-\rho\sigma^2 t^2/2)(1+t^2/\alpha^2)^{-\rho/2}(1+t^2/\beta^2)^{-\rho/2}\sin\rho(\mu t + +\theta_1 + \theta_2)) \to (0,0)$$

and

$$(Var(\mathrm{Re}\phi_n(t)), Var(\mathrm{Im}\phi_n(t)))$$
$$= \left(\frac{1}{2n}[\mathrm{Re}\phi(2t) + 1 - 2(\mathrm{Re}\phi(t))^2], \frac{1}{2n}[1 - \mathrm{Re}\phi(2t) - 2(\mathrm{Im}\phi(t))^2]\right)$$
$$= (\frac{1}{2n}[\exp(-2\rho\sigma^2 t^2)(1+4t^2/\alpha^2)^{-\rho/2}(1+4t^2/\beta^2)^{-\rho/2}\cos\rho(2\mu t + \theta_1 + \theta_2) + 1$$
$$-2(\exp(-\rho\sigma^2 t^2/2)(1+t^2/\alpha^2)^{-\rho/2}(1+t^2/\beta^2)^{-\rho/2}\cos\rho(\mu t + \theta_1 + \theta_2))^2],$$
$$\frac{1}{2n}[1 - \exp(-2\rho\sigma^2 t^2)(1+4t^2/\alpha^2)^{-\rho/2}(1+4t^2/\beta^2)^{-\rho/2}\cos\rho(2\mu t + \theta_1 + \theta_2)$$
$$-2(\exp(-\rho\sigma^2 t^2/2)(1+t^2/\alpha^2)^{-\rho/2}(1+t^2/\beta^2)^{-\rho/2}\sin\rho(\mu t+\theta_1+\theta_2))^2]) \to (\frac{1}{2n}, \frac{1}{2n})$$

while as $t \to 0$, these vectors tend to $(1,0)$ and $(0,0)$, respectively, since

$$\mathrm{Re}\phi(0) = 1 \textit{ and } \mathrm{Im}\phi(0) = 0.$$

Also, from relations (3.1.4), (3.3.2) and (3.3.3) it follows that, as $|t| \to \infty$,

$$Cov(\mathrm{Re}\phi_n(t), \mathrm{Im}\phi_n(t)) = \frac{1}{2n}\left[\mathrm{Im}\phi(2t) - \mathrm{Im}\phi(0) - 2\mathrm{Re}\phi(t)\mathrm{Im}\phi(t)\right]$$
$$= \frac{1}{2n}[\exp(-2\rho\sigma^2 t^2)(1+4t^2/\alpha^2)^{-\rho/2}(1+4t^2/\beta^2)^{-\rho/2}\sin\rho(2\mu t + \theta_1 + \theta_2)$$
$$-2\exp(-\rho\sigma^2 t^2)(1+t^2/\alpha^2)^{-\rho}(1+t^2/\beta^2)^{-\rho}\cos\rho(\mu t+\theta_1+\theta_2)\sin\rho(\mu t+\theta_1+\theta_2)] \to 0.$$

Since $\mathrm{Re}\phi(0) = 1$, $\mathrm{Im}\phi(0) = 0$, and

$$Cov(\mathrm{Re}\phi_n(t), \mathrm{Im}\phi_n(t)) = \frac{1}{2n}\left[\mathrm{Im}\phi(2t) - \mathrm{Im}\phi(0) - 2\mathrm{Re}\phi(t)\mathrm{Im}\phi(t)\right],$$

we have

$$Cov(\mathrm{Re}\phi_n(t), \mathrm{Im}\phi_n(t)) \longrightarrow 0 \ as \ t \longrightarrow 0.$$

Therefore, the tails of the two time series $\mathrm{Re}\phi_n(t)$ and $\mathrm{Im}\phi_n(t)$ are uncorrelated and $(\mathrm{Re}\phi_n(t), \mathrm{Im}\phi_n(t))$ estimates $(\mathrm{Re}\phi(t), \mathrm{Im}\phi(t))$ with increasing accuracy as $t \longrightarrow 0$.

## 3.4. Robustness of the quadratic distance estimator based on the empirical characteristic function

In the case of the quadratic distance estimator based on the empirical characteristic function, the functions $h_j(x)$, $j = 1, 2, ...., 2k$, as presented in remark 3.2.1., are defined as

$$h_j(x) = \cos(t_j x), \quad for \ \ j = 1, 2, ...., k$$

and

$$h_j(x) = \sin(t_{j-k} x), \quad for \ \ j = k+1, ...., 2k,$$

where $t_1, t_2, ..., t_k$ are fixed real numbers.
Since $|\cos x| \leq 1$ and $|\sin x| \leq 1$ for any real number $x$, we have that the functions $h_j(x)$, $j = 1, 2, ...., 2k$ are bounded functions of $x$ and by remark 2.3.1., it follows that this particular quadratic distance estimator is robust.

## 3.5. Hypothesis testing

### 3.5.1. Simple hypothesis

It is natural to use the quadratic distance to construct test statistics for testing the simple null hypothesis $\mathcal{H}_0$: $\mathbf{X}_1, \mathbf{X}_2, ...., \mathbf{X}_n$ come from a specified generalized normal Laplace (GNL) distribution with characteristic function

$$\phi_{\mathbf{X}}(t) = \left[\frac{\alpha_0\beta_0 \exp(\mu_0 it - \sigma_0^2 t^2/2)}{(\alpha_0 - it)(\beta_0 + it)}\right]^{\rho_0}.$$

Since $\theta_0 = [\mu_0, \sigma_0^2, \alpha_0, \beta_0, \rho_0]' = [\theta_1^0, \theta_2^0, \theta_3^0, \theta_4^0, \theta_5^0]'$ is specified, the test statistic

$$2nd(\theta_0) = 2n\,[\mathbf{Z}_n - \mathbf{Z}(\theta_0)]'\,\Sigma^{-1}(\theta_0)\,[\mathbf{Z}_n - \mathbf{Z}(\theta_0)]$$

can be considered. The values $t_1, t_2, ..., t_k$ are suitably chosen in a region of $t$ near zero, and are such that $\Sigma(\theta_0)$ is non-singular.

According to the results presented in subsection 2.4.2., we have that $2nd(\theta_0) \xrightarrow{\mathcal{L}} \chi^2_{2k}$, that is a chi-square test based on the value of the test statistic $2nd(\theta_0)$ can be performed.

In order to test the hypothesis $\mathcal{H}_0$ at significance level $\alpha$ (usually 5%), we compute the value of the test statistic $2nd(\theta_0)$ from the sample. If it is larger than the critical value $\chi^2_{2k,1-\alpha}$ (the $1-\alpha$ quantile of a $\chi^2$ distribution with $2k$ degrees of freedom), the null hypothesis $\mathcal{H}_0$ should be rejected.

### 3.5.2. Composite hypothesis

In constructing the test for the composite hypothesis
$\mathcal{H}_0$: $\mathbf{X}_1, \mathbf{X}_2, ...., \mathbf{X}_n$ come from a generalized normal Laplace (GNL) distribution with the characteristic function

$$\phi_{\mathbf{X}}(t) = \left[\frac{\alpha\beta \exp(\mu it - \sigma^2 t^2/2)}{(\alpha - it)(\beta + it)}\right]^{\rho}$$

$\alpha, \beta, \rho$ and $\sigma$ are positive parameters , $-\infty < \mu < +\infty$, $t$ is a real number, where the values of the parameters are not specified, we should first calculate the quadratic distance estimator $\widehat{\theta}^*$ by minimizing

$$d(\theta) = [\mathbf{Z}_n - \mathbf{Z}(\theta)]' \, \Sigma^{-1}(\theta) \, [\mathbf{Z}_n - \mathbf{Z}(\theta)]$$

with respect to $\theta$, or an equivalent expression where $\Sigma^{-1}(\theta)$ is replaced by a consistent estimate.

From theorem 2.4.2., given in subsection 2.4.3., it follows that the test statistic

$$2nd(\widehat{\theta}^*) = 2n \left[\mathbf{Z}_n - \mathbf{Z}(\widehat{\theta}^*)\right]' \Sigma^{-1}(\widehat{\theta}^*) \left[\mathbf{Z}_n - \mathbf{Z}(\widehat{\theta}^*)\right],$$

where $\Sigma^{-1}(\widehat{\theta}^*)$ can be replaced by another estimate of $\Sigma^{-1}(\theta)$ if desired, follows a limiting chi-square distribution $\chi^2_{2k-5}$ under $\mathcal{H}_0$ with $r = 2k - 5$ degrees of freedom.

Alternatively, since minimization of

$$d(\theta) = [\mathbf{Z}_n - \mathbf{Z}(\theta)]' \, \Sigma^{-1}(\theta) \, [\mathbf{Z}_n - \mathbf{Z}(\theta)]$$

involves the inverse of the matrix $\Sigma(\theta)$ which is dependent on the parameter, a simple procedure would be to replace $\Sigma(\theta)$ by the estimate $\Sigma_n$ defined analogously to $\Sigma(\theta)$ in terms of $\mathrm{Re}\phi_n(t)$ and $\mathrm{Im}\phi_n(t)$ and take as an estimator $\widetilde{\theta}^*$ the statistic minimizing

$$d'(\theta) = [\mathbf{Z}_n - \mathbf{Z}(\theta)]' \, \Sigma_n^{-1} \, [\mathbf{Z}_n - \mathbf{Z}(\theta)] \, .$$

The appropriate choice for a test statistic would then be the quadratic form

$$2nd'(\widetilde{\theta}^*) = 2n \left[\mathbf{Z}_n - \mathbf{Z}(\widetilde{\theta}^*)\right]' \Sigma_n^{-1} \left[\mathbf{Z}_n - \mathbf{Z}(\widetilde{\theta}^*)\right].$$

The consistency of $\phi_n(t)$ for $\phi(t)$ over finite intervals (result presented in section 3.1.) gives the consistency of $\Sigma_n$ for $\Sigma(\theta)$ under $\mathcal{H}_0$. Koutrouvelis and Kellermeier (1981) showed, based on this result, that $\widehat{\theta}^*$ and $\widetilde{\theta}^*$ are asymptotically equivalent estimators for $\theta$.

In order to test the hypothesis $\mathcal{H}_0$ at significance level $\alpha$ (usually 5%), we compute the value of the test statistic $2nd(\widehat{\theta}^*)$. If it is larger than the critical value $\chi^2_{2k-5,1-\alpha}$ (the $1 - \alpha$ quantile of a $\chi^2$ distribution with $2k - 5$ degrees of freedom), the hypothesis $\mathcal{H}_0$ should be rejected.

# Chapter 4

## NUMERICAL ILLUSTRATIONS

### 4.1. Simulation results

In this chapter we apply the quadratic distance method based on empirical characteristic function for estimation of the parameters of the generalized normal Laplace (GNL) distribution.

In section 1.2, we proved that a random variable $\mathbf{X}$ which follows the GNL distribution can be represented as:

$$\mathbf{X} \stackrel{d}{=} \rho\mu + \sigma\sqrt{\rho}\,\mathbf{Y} + (1/\alpha)\mathbf{G}_1 - (1/\beta)\mathbf{G}_2$$

where $\mathbf{Y}$, $\mathbf{G}_1$ and $\mathbf{G}_2$ are independent random variables with $\mathbf{Y} \sim \mathbf{N}(0,\ 1)$, i.e. with probability density function $f(x) = \frac{1}{\sqrt{2\pi}} \exp(-x^2/2)$, $x \in \mathbf{R}$ and $\mathbf{G}_1$, $\mathbf{G}_2$ gamma random variables with shape parameter $\rho$ and scale parameter 1, i.e. with probability density function $g(x) = \frac{1}{\Gamma(\rho)} x^{\rho-1} \exp(-x)$, $x > 0$.

This representation provides a straightforward way to generate observations from a GNL distribution.

We started by generating samples of 500, 1000 and 10000 GNL random variables with parameters

$$\mu = 3,\ \sigma^2 = 4,\ \alpha = 2,\ \beta = 4 \textit{ and } \rho = 4.$$

We calculated the QDE

$$\widehat{\theta} = [\widehat{\mu},\ \widehat{\sigma}^2,\ \widehat{\alpha},\ \widehat{\beta},\ \widehat{\rho}\,]$$

with various choices for the matrix $Q(\theta)$, the points $t_1, t_2, \ldots, t_k$ and the number of points $k$. First, we obtained the estimates of the QDE with the matrix $Q(\theta)$

equal to the identity matrix $I$. The estimates are obtained by minimizing the distance

$$d(\theta) = [\mathbf{Z}_n - \mathbf{Z}(\theta)]'[\mathbf{Z}_n - \mathbf{Z}(\theta)],$$

that is

$$d(\theta) = \sum_{j=1}^{k} \left[ (\mathrm{Re}\phi_n(t_j) - \mathrm{Re}\phi(t_j))^2 + (\mathrm{Im}\phi_n(t_j) - \mathrm{Im}\phi(t_j))^2 \right].$$

We considered different sets of points for $t_j$, $j = 1, 2, \ldots, k$ and for $k$ and we concluded which are the best choices we should consider in the process of the estimation. The situations we considered in choosing the points $t_j$, $j = 1, 2, \ldots, k$ are: from 0.1 to 1, 2 and respectively 5 in step size of 0.1 (situations A, B and C), from 0.01 to 1 in step size of 0.01 (situation D), from 0.05 to 1 in step size of 0.05 (situation E) and from 0.001 to 0.02 in step size of 0.001 (situation F) and from 0.0001 to 0.002 in step size of 0.0001 (situation G).

For calculating the real and imaginary parts of the empirical characteristic function, $\phi_n(t)$, and theoretical one, $\phi(t)$, we used the functions ComplexExpand[Re] and ComplexExpand[Im] from MATHEMATICA. With the FindMinimum procedure in MATHEMATICA, the estimates of the QDE were easily obtained.

TABLE 4.1. Estimated values of the QDE with $Q(\theta) = I$ using a sample of size 500

| $t_1, t_2, \ldots, t_k$ | $\widehat{\mu}$ | $\widehat{\sigma}^2$ | $\widehat{\alpha}$ | $\widehat{\beta}$ | $\widehat{\rho}$ |
|---|---|---|---|---|---|
| A: 0.1, 0.2, ...,1 | 3.47199 | 1.79137 | 0.744489 | 0.836748 | 3.64204 |
| B: 0.1, 0.2, ...,2 | 3.03235 | 3.62886 | 0.973442 | 2.59586 | 3.59742 |
| C: 0.1, 0.2, ...,5 | 3.03126 | 3.62754 | 0.974811 | 2.608922 | 3.61742 |
| D: 0.01, 0.02,...,1 | 3.46325 | 2.10581 | 0.784382 | 0.897941 | 3.63723 |
| E: 0.05,0.1,0.15,...,1 | 3.4671 | 2.01312 | 0.772177 | 0.878252 | 3.63826 |
| F: 0.001,0.002,...,0.02 | 3.05092 | 3.7089 | 2.31208 | 3.8066 | 4.0528 |
| G: 0.0001,0.0002,...,0.002 | 2.90725 | 4.13327 | 2.07241 | 3.95164 | 4.10989 |

TABLE 4.2. Estimated values of the QDE with $Q(\theta) = I$ using a sample of size 1000

| $t_1, t_2, \ldots, t_k$ | $\widehat{\mu}$ | $\widehat{\sigma}^2$ | $\widehat{\alpha}$ | $\widehat{\beta}$ | $\widehat{\rho}$ |
|---|---|---|---|---|---|
| A: 0.1, 0.2, ...,1 | 2.33522 | 1.37416 | 1.0844 | 1.17815 | 4.94689 |
| B: 0.1, 0.2, ...,2 | 3.74402 | 2.64488 | 1.94689 | 0.971797 | 4.52567 |
| C: 0.1, 0.2, ...,5 | 2.99224 | 1.67399 | 0.849408 | 0.983071 | 4.11173 |
| D: 0.01, 0.02,...,1 | 2.11573 | 0.986332 | 0.957653 | 1.09653 | 5.76391 |
| E: 0.05,0.1,0.15,...,1 | 2.06891 | 0.938899 | 0.961799 | 1.0999 | 5.89114 |
| F: 0.001,0.002,...,0.02 | 2.9794 | 3.85397 | 1.71799 | 4.01907 | 3.98061 |
| G: 0.0001,0.0002,...,0.002 | 2.96931 | 4.25337 | 1.99308 | 4.00491 | 4.05235 |

TABLE 4.3. Estimated values of the QDE with $Q(\theta) = I$ using a sample of size 10000

| $t_1, t_2, \ldots, t_k$ | $\widehat{\mu}$ | $\widehat{\sigma}^2$ | $\widehat{\alpha}$ | $\widehat{\beta}$ | $\widehat{\rho}$ |
|---|---|---|---|---|---|
| A: 0.1, 0.2, ...,1 | 2.49034 | 1.76831 | 1.0089 | 1.13663 | 4.99584 |
| B: 0.1, 0.2, ...,2 | 2.95635 | 2.22516 | 0.956776 | 1.08182 | 4.22423 |
| C: 0.1, 0.2, ...,5 | 2.94721 | 2.2571 | 0.978133 | 1.10225 | 4.223212 |
| D: 0.01, 0.02,...,1 | 2.52268 | 1.81036 | 1.00902 | 1.1363 | 4.93508 |
| E: 0.05,0.1,0.15,...,1 | 2.51578 | 1.80174 | 1.00822 | 1.1367 | 4.94622 |
| F: 0.001,0.002,...,0.02 | 2.96532 | 3.92057 | 1.7794 | 3.95461 | 3.96226 |
| G: 0.0001,0.0002,...,0.002 | 3.02248 | 4.02163 | 2.22251 | 4.00138 | 4.02332 |

Tables 4.1, 4.2 and 4.3 contain the values considered for the points $t_j$ and the estimates of the QDE obtained using the generated samples of 500, 1000 and 10000 observations, respectively.

Tables 4.4, 4.5 and 4.6 contain the absolute biases of the quadratic distance estimators obtained in the situations mentioned above.

TABLE 4.4. Absolute biases of the QDE with $Q(\theta) = I$ using a sample of size 500

| $t_1, t_2, \ldots, t_k$ | $\widehat{\mu}$ | $\widehat{\sigma}^2$ | $\widehat{\alpha}$ | $\widehat{\beta}$ | $\widehat{\rho}$ |
|---|---|---|---|---|---|
| A: 0.1, 0.2, ...,1 | 0.47199 | 2.20863 | 1.255511 | 3.163252 | 0.35796 |
| B: 0.1, 0.2, ...,2 | 0.03235 | 0.37114 | 1.026558 | 1.40414 | 0.40258 |
| C: 0.1, 0.2, ...,5 | 0.03126 | 0.37246 | 1.025189 | 1.391078 | 0.38258 |
| D: 0.01, 0.02,...,1 | 0.46325 | 1.89419 | 1.215618 | 3.102059 | 0.36277 |
| E: 0.05,0.1,0.15,...,1 | 0.4671 | 1.98688 | 1.227823 | 3.121748 | 0.36174 |
| F: 0.001,0.002,...,0.02 | 0.05092 | 0.2911 | 0.31208 | 0.1934 | 0.0528 |
| G: 0.0001,0.0002,...,0.002 | 0.09275 | 0.13327 | 0.07241 | 0.04836 | 0.10989 |

TABLE 4.5. Absolute biases of the QDE with $Q(\theta) = I$ using a sample of size 1000

| $t_1, t_2, \ldots, t_k$ | $\widehat{\mu}$ | $\widehat{\sigma}^2$ | $\widehat{\alpha}$ | $\widehat{\beta}$ | $\widehat{\rho}$ |
|---|---|---|---|---|---|
| A: 0.1, 0.2, ...,1 | 0.66478 | 2.62584 | 0.9156 | 2.82185 | 0.94689 |
| B: 0.1, 0.2, ...,2 | 0.74402 | 1.35512 | 0.05311 | 3.028203 | 0.52567 |
| C: 0.1, 0.2, ...,5 | 0.00776 | 2.32601 | 1.150592 | 3.016929 | 0.11173 |
| D: 0.01, 0.02,...,1 | 0.88427 | 3.013668 | 1.042347 | 2.90347 | 1.76391 |
| E: 0.05,0.1,0.15,...,1 | 0.93109 | 3.061101 | 1.038201 | 2.9001 | 1.89114 |
| F: 0.001,0.002,...,0.02 | 0.0206 | 0.14603 | 0.28201 | 0.01907 | 0.01939 |
| G: 0.0001,0.0002,...,0.002 | 0.03069 | 0.25337 | 0.00692 | 0.00491 | 0.05235 |

Based on these results, we notice that the estimates of the QDE in the situations A, B, C, D and E have large bias for some parameters and a bit smaller for others, regardless the sample size. We conclude that large values of $t_j$ produce worst estimates for $\widehat{\sigma}^2$, $\widehat{\alpha}$ and $\widehat{\beta}$.

For values very small (near zero) like those from the situations F and G, we obtained good estimates of the QDE with small bias for all the five parameters and are better than the results obtained in the situations A, B, C, D and E.

TABLE 4.6. Absolute biases of the QDE with $Q(\theta) = I$ using a sample of size 10000

| $t_1, t_2, \ldots, t_k$ | $\widehat{\mu}$ | $\widehat{\sigma}^2$ | $\widehat{\alpha}$ | $\widehat{\beta}$ | $\widehat{\rho}$ |
|---|---|---|---|---|---|
| A: 0.1, 0.2, ...,1 | 0.50966 | 2.23169 | 0.9911 | 2.86337 | 0.99584 |
| B: 0.1, 0.2, ...,2 | 0.04365 | 1.77484 | 1.043224 | 2.91818 | 0.22423 |
| C: 0.1, 0.2, ...,5 | 0.05279 | 1.7429 | 1.021867 | 2.89775 | 0.23212 |
| D: 0.01, 0.02,...,1 | 0.47732 | 2.18964 | 0.99098 | 2.8637 | 0.93508 |
| E: 0.05,0.1,0.15,...,1 | 0.48422 | 2.19826 | 0.99178 | 2.8633 | 0.94622 |
| F: 0.001,0.002,...,0.02 | 0.03468 | 0.07943 | 0.2206 | 0.04539 | 0.03774 |
| G: 0.0001,0.0002,...,0.002 | 0.02248 | 0.02163 | 0.22251 | 0.00138 | 0.02332 |

Next, we focused on the estimator $\widehat{\theta^\star} = [\widehat{\mu^\star},\ \widehat{\sigma^\star}^2,\widehat{\alpha^\star},\ \widehat{\beta^\star},\ \widehat{\rho^\star}\ ]$ obtained by minimizing the distance

$$d(\theta) = [\mathbf{Z}_n - \mathbf{Z}(\theta)]'Q(\theta)[\mathbf{Z}_n - \mathbf{Z}(\theta)].$$

First, we tried with the optimal choice $Q(\theta) = \Sigma^{-1}(\theta)$. Since in the situations F and G the number of points $t_j$ is 20, the dimensions of the corresponding matrices $\Sigma(\theta)$ are $40 \times 40$ and we had difficulties in calculating symbolically their inverses. So, instead of $Q(\theta)$ we used the consistent estimator $\widehat{Q}(\theta) = \Sigma^{-1}(\widehat{\theta})$ with the values of $\widehat{\theta} = [\widehat{\mu},\ \widehat{\sigma}^2,\ \widehat{\alpha},\ \widehat{\beta},\ \widehat{\rho}\,]$ obtained in the situations F and G (the values are found in the last two lines of tables 4.1, 4.2, 4.3). We obtained the estimates of the QDE $\widehat{\theta^\star} = [\widehat{\mu^\star},\ \widehat{\sigma^\star}^2,\widehat{\alpha^\star},\ \widehat{\beta^\star},\ \widehat{\rho^\star}\ ]$ by minimizing the distance

$$d(\theta) = [\mathbf{Z}_n - \mathbf{Z}(\theta)]'\widehat{Q}(\theta)[\mathbf{Z}_n - \mathbf{Z}(\theta)],$$

where

$$[\mathbf{Z}_n - \mathbf{Z}(\theta)] = [\mathrm{Re}\phi_n(t_1) - \mathrm{Re}\phi(t_1), \ldots, \mathrm{Im}\phi_n(t_{20}) - \mathrm{Im}\phi(t_{20})]'$$

and the matrix $\Sigma(\widehat{\theta})$ is a $40 \times 40$ matrix with the elements $\left(\sigma_{i,j}(\widehat{\theta})\right)_{1\le i,j\le 40}$ given in section 3.2, that is

$$\sigma_{i,j}(\widehat{\theta}) = \begin{cases} \frac{1}{2}[\mathrm{Re}\phi(t_i+t_j) + \mathrm{Re}\phi(t_i-t_j)] - \mathrm{Re}\phi(t_i)\mathrm{Re}\phi(t_j) & \text{for } 1 \le i,\ j \le 20 \\ \frac{1}{2}[\mathrm{Re}\phi(t_i-t_j) - \mathrm{Re}\phi(t_i+t_j)] - \mathrm{Im}\phi(t_i)\mathrm{Im}\phi(t_j) & \text{for } 21 \le i,\ j \le 40 \\ \frac{1}{2}[\mathrm{Im}\phi(t_i+t_j) - \mathrm{Im}\phi(t_i-t_j)] - \mathrm{Re}\phi(t_i)\mathrm{Im}\phi(t_j) & \text{for } 1 \le i \le 20,\ 21 \le j \le 40, \end{cases}$$

where we define $t_i = t_{i-20}\ for\ 21 \leq i \leq 40$. Tables 4.7, 4.8 and 4.9 contain the estimates of the estimator $\widehat{\theta}^*$ in the situations F and G.

TABLE 4.7. Estimated values of the QDE with $Q(\theta) = \Sigma^{-1}(\widehat{\theta})$ (size 500)

| $t_1, t_2, \ldots, t_k$ | $\widehat{\mu}^*$ | $\widehat{\sigma^*}^2$ | $\widehat{\alpha}^*$ | $\widehat{\beta}^*$ | $\widehat{\rho}^*$ |
|---|---|---|---|---|---|
| F: 0.001,0.002,...,0.02 | 2.98649 | 4.00359 | 1.99249 | 3.97338 | 3.99148 |
| G: 0.0001,0.0002,...,0.002 | 2.9889 | 3.99875 | 1.99897 | 4.000582 | 3.99795 |

TABLE 4.8. Estimated values of the QDE with $Q(\theta) = \Sigma^{-1}(\widehat{\theta})$ (size 1000)

| $t_1, t_2, \ldots, t_k$ | $\widehat{\mu}^*$ | $\widehat{\sigma^*}^2$ | $\widehat{\alpha}^*$ | $\widehat{\beta}^*$ | $\widehat{\rho}^*$ |
|---|---|---|---|---|---|
| F: 0.001,0.002,...,0.02 | 3.01002 | 4.00073 | 2.00073 | 3.98998 | 3.99699 |
| G: 0.0001,0.0002,...,0.002 | 3.00205 | 4.00035 | 2.00052 | 4.00374 | 3.99882 |

TABLE 4.9. Estimated values of the QDE with $Q(\theta) = \Sigma^{-1}(\widehat{\theta})$ (size 10000)

| $t_1, t_2, \ldots, t_k$ | $\widehat{\mu}^*$ | $\widehat{\sigma^*}^2$ | $\widehat{\alpha}^*$ | $\widehat{\beta}^*$ | $\widehat{\rho}^*$ |
|---|---|---|---|---|---|
| F: 0.001,0.002,...,0.02 | 2.99783 | 4.00059 | 1.99934 | 3.99894 | 3.99873 |
| G: 0.0001,0.0002,...,0.002 | 3.00027 | 4.00004 | 2.000241 | 4.00011 | 3.99881 |

In all of these 6 cases, after these two iterations the algorithm converged. Thus the estimators are consistent.

The following tables contain the estimated standard deviations of the estimators $\widehat{\mu}^*$, $\widehat{\sigma^*}^2$, $\widehat{\alpha}^*$, $\widehat{\beta}^*$ and $\widehat{\rho}^*$ calculated from formula $\widehat{\theta}^* = (1/n)(S'\Sigma^{-1}S)^{-1}$, where $S = (\frac{\partial \mathbf{Z}_i(\theta)}{\partial \theta_j})$ is a matrix of dimension $40 \times 5$, with $i = 1, 2, ...., 40$ $j = 1, 2, 3, 4, 5$, and $\mathbf{Z}_i(\theta) = \begin{cases} \mathrm{Re}\phi(t_i) & \text{for } i = 1, 2, \ldots, 20 \\ \mathrm{Im}\phi(t_{i-20}) & \text{for } i = 21, \ldots, 40. \end{cases}$

TABLE 4.10. Estimated standard deviation (size 500)

| *situations* | $sd(\widehat{\mu}^{\star})$ | $sd(\widehat{\sigma^{\star}}^2)$ | $sd(\widehat{\alpha}^{\star})$ | $sd(\widehat{\beta}^{\star})$ | $sd(\widehat{\rho}^{\star})$ |
|---:|---:|---:|---:|---:|---:|
| F | 0.021534 | 0.00366 | 0.00546 | 0.00135 | 0.012707 |
| G | 0.020451 | 0.00127 | 0.00538 | 0.00124 | 0.011909 |

TABLE 4.11. Estimated standard deviation (size 1000)

| *situations* | $sd(\widehat{\mu}^{\star})$ | $sd(\widehat{\sigma^{\star}}^2)$ | $sd(\widehat{\alpha}^{\star})$ | $sd(\widehat{\beta}^{\star})$ | $sd(\widehat{\rho}^{\star})$ |
|---:|---:|---:|---:|---:|---:|
| F | 0.01522667 | 0.002589826 | 0.00385761 | 0.00095597 | 0.0126903 |
| G | 0.01522718 | 0.002589498 | 0.00385547 | 0.00095406 | 0.0126682 |

TABLE 4.12. Estimated standard deviation (size 10000)

| *situations* | $sd(\widehat{\mu}^{\star})$ | $sd(\widehat{\sigma^{\star}}^2)$ | $sd(\widehat{\alpha}^{\star})$ | $sd(\widehat{\beta}^{\star})$ | $sd(\widehat{\rho}^{\star})$ |
|---:|---:|---:|---:|---:|---:|
| F | 0.00481369 | 0.00081882 | 0.0012220 | 0.00030235 | 0.004006 |
| G | 0.00481211 | 0.00081701 | 0.0012168 | 0.00030224 | 0.004004 |

The obtained estimates of the QDE in all of these cases have small biases and standard deviations.

We also generated samples of GNL random variables for other values of the parameters and obtained good estimates.

# Chapter 5

## CONCLUSION

In this thesis, we developed a method to estimate the parameters of the generalized normal Laplace distribution, by minimizing the distance between the real and imaginary parts of the theoretical characteristic functions and their empirical counterparts.

We obtained an expression for the variance-covariance matrix of the terms of errors and we constructed the estimators. The estimators obtained are robust, consistent and have an asymptotic normal distribution, from which confidence intervals can be constructed.

We constructed test statistics based on the quadratic distance for testing the goodness-of-fit of the generalized normal Laplace distribution for simple and composite hypothesis. For model testing, the tests statistics follow a chi-square distribution, which makes the tests statistics easy to use.

We simulated several samples of data sets following the generalized normal Laplace distribution and we obtained very good estimators.

Quadratic distance methods are fairly versatile and easy to implement and can be considered as alternative methods to maximum likelihood estimation in cases when the likelihood function is difficult to compute, to the method of moments and to the minimum chi-square methods.

Luong and Doray (2002) showed that in the case of discrete distributions, the quadratic distance estimator is flexible, offering a trade-off between efficiency and robustness.

# BIBLIOGRAPHY


[1] EPPS T.W. AND PULLEY L.B., *Parameter estimates and tests of fit for infinite mixture distributions*, Comm. Statist.-Theory and Methods, **14**, 3125-3145, 1985.

[2] FELLER W., *An introduction to probability theory and its applications*, Vol.2, John Wiley and Sons, New York, 1971.

[3] FEUERVERGER A. AND MCDUNNOUGH P., *On the efficiency of empirical characteristic function procedures*, J. Roy. Statist. Soc., Ser. B, **43**, 20-27, 1981.

[4] FEUERVERGER A. AND MUREIKA R.A., *The empirical characteristic function and its applications*, The Annals of Statistics, **1**, 88-97, 1977.

[5] HAHN J., *LaTeX for Everyone: A reference guide and tutorial for typesetting documents using a computer*, Personal TeX, Inc., Mill Valley, CA., 1991.

[6] HAMPEL F.R., *Contributions to the theory of robust etimation*, Ph.D.Thesis, Dep. of Statistics, Univ. of California, Berkeley, 1968.

[7] HAMPEL F.R., *The influence curve and its role in robust estimation*, J. Amer. Statist. Assoc., **69**, 383-393, 1974.

[8] HUBERMAN B. AND ADAMIC L.A., *Internet: growth dynamics of the World-Wide web.*, Nature, **401**, 130-131, 1999.

[9] KENDALL M.G. AND STUART A., *The advanced theory of statistics, Vol. 1*, Charles Griffin and Sons, London, 1969.

[10] KOTZ S. AND KOZUBOWSKI T.J. AND PODGÓRSKI K., *The Laplace distribution and generalizations*, Birkhäuser, Boston, 2001.

[11] KOUTROUVELIS I.A. AND KELLERMEIER J., *A goodness of fit test based on the empirical characteristic function when parameters must be estimated*, J. Roy. Statist. Soc., **43**, 173-176, 1981.

[12] LUONG A. AND DORAY L.G., *General quadratic distance methods for discrete distributions definable recursively*, Insurance: Mathematics and Economics, **30**, 255-267, 2002.

[13] LUONG A. AND DORAY L.G., *Inference for the positive stable laws based on a special quadratic distance*, submitted to North American Actuarial Journal, 2006.

[14] LUONG A. AND THOMPSON M.E., *Minimum-distance methods based on quadratic distances for transforms*, The Canadian Journal of Statistics, **3**, 239-251, 1987.

[15] MITZENMACHER M., *Dynamic models for file sizes and double Pareto distributions*, Draft manuscript available at http://www.eecs.harvard.edu/michaelm/NEWWORK/papers/, 2001.

[16] MOORE D.S., *Generalized inverses, Wald's method and the construction of chi-square tests of fit*, J. Amer. Statist. Assoc., **72**, 131-137, 1977.

[17] MOORE D.S., *Chi-square tests*, Studies in Statistics, Vol. 19, Mathematical Association of America, 1978.

[18] PAULSON A.S. AND HALCOMB E.W. AND LEITCH R.A., *The estimation of the parameters of the stable laws*, Biometrika, **62**, 163-170, 1975.

[19] RAO C.R., *Linear statistical inference and its applications*, Wiley, New York, 1973.

[20] REED W.J., *The normal-Laplace distribution and its relatives*, to appear Birkhäuser, 2004.

[21] REED W.J., *Brownian-Laplace Motion and its use in financial modelling*, Commun. Stat. A., 2005.

[22] REED W.J. AND JORGENSEN M., *The double Pareto-lognormal distribution-A new parametric model for size distributions*, Commun. Stat.-Theory and Methods, **33**, 1733-1753, 2004.

[23] RYDBERG T.H., *Realistic statistical modelling of financial data*, Inter. Stat. Rev., **68**, 233-258, 2000.

[24] SCHOUTENS W., *Lévy processes in finance*, John Wiley and Sons, Chichester, 2003.

[25] SHIRYAYEV A.N., *Probability*, Springer-Verlag New York, 1984.

[26] SPIVAK M.D., *The Joy of TEX*, second edition, Amer. Math. Soc., Providence, R.I., 1990.

[27] TEX USER GROUP, *LATEX2e for classand package writers*, disponible sur le web, 1995.