



HAL
open science

The Lila distribution and its applications in risk modelling

Bertrand K. Hassani, Wei Yang

► **To cite this version:**

Bertrand K. Hassani, Wei Yang. The Lila distribution and its applications in risk modelling. 2016. halshs-01400186

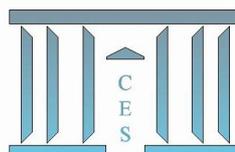
HAL Id: halshs-01400186

<https://shs.hal.science/halshs-01400186>

Submitted on 21 Nov 2016

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



**The Lila distribution and its applications
in risk modelling**

Bertrand HASSANI, Wei YANG

2016.68



The Lila distribution and its applications in risk modelling

Bertrand Hassani ¹ and Wei Yang ²

Disclaimer: The opinions, ideas and approaches expressed or presented are those of the authors and do not necessarily reflect Santander's position. As a result, Santander cannot be held responsible for them.

Abstract: Risk data sets tend to have heavy-tailed, sometimes bi-modal, empirical distributions, especially in operational risk, market risk and customers behaviour data sets. To capture these observed "unusual" features, we construct a new probability distribution and call it the lowered-inside-leveraged-aside (Lila) distribution as it transfers the embedded weight of data from the body to the tail. This newly constructed distribution can be viewed as a parametric distribution with two peaks. It is constructed through the composition of a Sigmoid-shaped continuous increasing differentiable function with cumulative distribution functions of random variables. Examples and some basic properties of the Lila distribution are illustrated. As an application, we fit a Lila distribution to a set of generated data by using the quantile distance minimisation method (alternative methodologies have been tested too, such as maximum likelihood estimation).

Key words: probability distribution, parametric distribution, multimodal distribution, operational risk, market risk, pseudo bi-modal distribution.

1 Introduction

Since the Dodd-Frank Act was signed in 2010, there has been a wave of new regulations to reform the financial industry and to improve the stability of financial system (coming from the Bank of England, the European Banking Authority, the Basel Committee, the Federal Reserve, etc.). In order to

¹Grupo Santander and Universit Paris 1 Panthéon-Sorbonne CES UMR 8174, 106 boulevard de l'Hopital 75647 Paris Cedex 13, France, +44 (0)2070860973, bertrand.hassani@gruposantander.com

²Risk methodology, Santander UK plc, 2 Triton Square, Regent's Place, London NW1 3AN, +44 (0)2077565193, wei.yang@santander.co.uk

meet formal expectations set by regulators, banks have elevated standards for governance and enterprise risk management. Internal risk reporting practices are required to be set up and risk data are aggregated according to supervisory recommendations. Standard models or advanced internal models approved by regulators are used to calculate the amount of capital which should be reserved, this type of capital is usually referred to as regulatory capital. Models relying on statistical distributions are also used to evaluate the economic capital and more importantly these models are used for management purposes, i.e. to make decisions.

Indeed, any model for regulatory capital calculation requires a way of quantifying the level of risk. In traditional statistics, a commonly used risk measure is the standard deviation. Yet within the context of risk management and capital requirements, the most common measure of risk is a quantile-based risk measure Value-at-Risk (*VaR*). The VaR measures the maximum potential loss with a degree of confidence for a specified period. Given a confidence level $\alpha \in (0, 1)$, the VaR_α is the relevant quantile of the loss distribution:

$$VaR_\alpha(X) = \inf\{x | P[X > x] \leq 1 - \alpha\} = \inf\{x | F_X(x) \geq \alpha\}$$

where X is a risk factor admitting a loss distribution F_X , for a systematic discussion on *VaR* see e.g. [10]. Another choice of risk measure as an extension of the *VaR* is the Expected Shortfall (*ES*, also called Conditional VaR), which seeks to assess what happens beyond its maximum loss threshold. Specifically, ES_α is defined as the average of all losses which are greater than or equal to VaR_α c.f. [22]:

$$ES_\alpha(X) = \frac{1}{1 - \alpha} \int_\alpha^1 VaR_p dp.$$

While there remain a number of issues related to the selection of appropriate risk measures since every risk measure has drawbacks and limited applications, a prerequisite for any risk model is the modelling of the underlying risks from historical data. There is a wide range of distribution families available for risk modelling since many distributions, such as log-normal, Weibull, Gamma, and Generalised Hypobolic [3] distributions, are successfully used to model real life random variables. However, irrelevant of risk types, most of these uni-modal parametric distributions fail to capture fat-tailed features of risk data, in the sense that upper quantiles are larger than those obtained from any fitted distributions. In order to model

heavy-tailed risk data, extreme value distribution and heavy-tailed pareto-type distributions with a shape parameter β have attracted a lot attention. The tails of the loss distribution functions are in the form:

$$\mathbb{P}(X > x) = x^{-\beta}h(x)$$

where $\beta > 0$ is the tail-index parameter and h is a slowly varying function; for detailed discussion see e.g. [8]. Yet for example, the risk measure VaR , which is currently the financial industry practice, is sensitive to the value of β , especially when $0 < \beta < 1$. In case $0 < \beta < 1$, that is the distribution has a extremely heavy tail, the VaR may lead to ridiculously high capital charges, besides the fact that VaR no longer has the desired property of sub-additivity [21]. (The lack of sub-additivity of VaR contradicts the diversification effect.) Even in some cases, tails of risk data are too "fat" such that another "peak" is formed.

Remark 1.1. *Though in this paper we are focusing on risk measures in the tail, by properly fitting the data, the modelling of the expected loss (represented by the mean or the median for example depending on how financial institutions perceive it) is also impacted, consequently the risk management decision might be significantly modified.*

The characteristics of fat-tails and even multiple peaks of risk data make it impossible to appropriately fit and use most uni-modal parametric distributions. To tackle these issues, mixture distributions have been proposed and applied in the area of risk modelling. The importance of mixture distributions is remarked in a number of books on mixtures for example [20, 9] and a diversity of publications. Mao [18] identified a variety of sources of difficulties in statistical inference using mixture.

The application of mixture distributions has also seen some drawbacks in the area of risk modelling. The mixture of Gaussian distributions to create so-called Gaussian Mixture (GM) distributions are considered e.g. in [5, 19]. GM distributions can accommodate heavy tails, asymmetry, and/or multimodality. Given enough components, the mixture of normal distributions can achieve great flexibility. For example, the mixture of two Normal distributions requires 5 parameters, 2 for each normal distribution and 1 weight parameter. However, the weight parameter in particular might be quite complicated to obtain. Besides, the support of the resultant distribution is infinite which makes it complicated to be applied to credit and operational risk, although this mixed distribution may be appropriate to model market

risk.

Guégan and Hassani [11] applied the mixture of a lognormal and a Generalized Pareto Distribution (GPD) for modelling operational risk and focused on the estimation of the threshold. Although the resulting distribution fits better operational risk data, the proposed method involves the estimation of 5 parameters. Particularly the estimation procedure for the truncated lognormal distribution, the estimation of the location parameter and the shape parameter of the GPD make it a complicated approach to undertake in practice.

One may consider to combine distributions of any kind. By this method, the number of parameters needed increases. Usually this method requires at least 5 parameters. The resulting distribution is very unstable as a slight change in the data set such as adding a new data point may require the modification of the parameters and/or new combination of distributions (a feature that multiple approved OpRisk models permit).

In this paper, we introduce a new class of probability distributions that features a parametric distribution with two peaks of different heights. The introduction is called the lowered-inside-leveraged-aside (Lila) distribution. The Lila distribution is inspired by the unusual phenomenon observed from the risk data. For example, in risk data sets, small losses have very high frequency and only a few extremely large losses appear. However, the extreme large losses are not negligible since the sum of the few large losses accounts for a large proportion of the total losses. Sometimes, the empirical distributions of loss data often exhibit two peaks which are of different heights and far apart from each other.

Informally, to construct the distribution with the desired features, the idea is to distort a cumulative distribution function of a random variable, by "re-distributing" the probabilities of events. We lower the probabilities of rare events in the body part and leverage those in the tail events to form a "tail peak". Mathematically speaking, a S-shaped function is used to assign weights to events and then it is composed with the distribution under consideration. Formally we call this new class of distributions *lowered-inside-leveraged-aside (Lila) distributions*. Due to the feature of two peaks with different heights, we refer to this new class of Lila distributions as *bi-modal parametric distributions*. With abuse of terminology, the term *modal* is used for local maxima of the Lila density function at which the density's derivative vanishes and the second order derivative is negative.

The rest of the paper is organised as follows: in Section 2, we formally construct the Lila distributions and discuss some basic properties. The method is illustrated through an example in Section 3, with detailed analysis on the roles played by the parameters used. In Section 4, we fit two sets of real data to the Lila distribution as an application. Finally Section 5 concludes the paper.

2 Construction and properties of Lila family of distributions

2.1 Construction of the Lila distribution

Let X be a non-negative random variable with a cumulative distribution function $F_X(x) = P(X \leq x)$ where $\lim_{x \rightarrow 0} F(x) = 0$ and $\lim_{x \rightarrow \infty} F(x) = 1$. Let F_X^{-1} denote the inverse function ³ of F_X and f_X the density function of F_X .

Consider a distortion function $G : [0, 1] \rightarrow [0, 1]$ with $G(0) = 0$ and $G(1) = 1$ which is continuous and increasing with a differentiable and convex derivative function $g : [0, 1] \rightarrow [0, k]$, for a $k > 0$. This type of functions G are often referred to as sigmoid functions or S -shape functions. In the paper, we study the Lila distributions with the specific polynomial G in Eq. (3.1) and Eq. (3.3).

Definition 2.1. Define a function $L : [0, \infty) \rightarrow [0, 1]$ by

$$L(y) := (G \circ F_X)(y) = G(F_X(y)), \quad \forall y \in [0, \infty) \quad (2.1)$$

where the notation \circ denotes the composition of two functions. Denote the density function of L by $l : [0, \infty) \rightarrow [0, \infty)$

$$l(x) := (G \circ F_X)'(x) = (g \circ F_X)(x)f_X(x). \quad (2.2)$$

The distribution function L defined in (2.1) with the density function l in (2.2) is called the lowered-inside-leveraged-aside (Lila) distribution.

Remark 2.1. By construction, the function L defined in (2.1) satisfies the criterion of being a cumulative distribution function. That is, the function L is a non-decreasing continuous function with $\lim_{x \rightarrow 0} L(x) = 0$ and $\lim_{x \rightarrow \infty} L(x) = 1$.

³Sometimes F_X^{-1} is referred to as quantile function of X .

Remark 2.2. *Effectively, one can understand that the construction of Lila distributions is to "re-distribute" the density of the underlying random variable according to the "weight" $(g \circ F_X)(x)$ for each x .*

We would like to stress that, in order to generate general Lila distributions, the distortion function G must satisfy the following conditions:

1. $0 < G(p) < 1$, $G(0) = 0$ and $G(1) = 1$. This condition ensures that for each x , $(G \circ F_X)(x)$ defines a valid probability.
2. $G(p)$ is a differentiable and increasing function, i.e. $g(p) \geq 0$. This condition ensures that the risk is re-distributed (or shifted) to the tail part and the distorted probability $(G \circ F_X)(x)$ defines another distribution.
3. The derivative function $g(p)$ is non-negative, differentiable and convex. This condition is critical and ensures the resultant distribution is smooth and has two peaks.

The above method to construct a new distribution function is related to the method of *transformation of random variables*, that is using an invertible transformation function $y(x)$ to transform a random variable X into another one Y . Then the density function of the new random variable Y can be implied from the transformation function. In contrast, our method directly distorts the distribution of the random variable X . One can find the transformation function $y(x)$, which transforms a random variable X with the distribution $f_X(x)$ into a new random variable Y with the Lila distribution, by solving the following equation for $y(x)$

$$g(F_X(y))f_X(y) = f_X(x(y)) \left| \frac{dx}{dy} \right|. \quad (2.3)$$

Another related method to construct new distributions is through the composition of distributions functions. For example, in [17], Jones chose Beta distributions for g with parameters a, b and constructed a family of distributions by specifying F_X to be normal, Fréchet, Gumbel, exponential Weibull, exponential, Burr and power distributions. Yet the Beta distribution g in [17] is either convex or concave depending on the parameters a, b . In [1], AL-Hussaini generated distribution functions by composing a cumulative distribution function F_X with several cumulative distribution functions G including Weibull, exponential, Rayleigh, Burr type XII, Lomax, compound Rayleigh, Gompertz and compound Gompertz CDFs. However,

AL-Hussaini focused on the study of the so-called survival function in the form $1 - G[-\ln F_X(y)]$.

Our method is also different from constructing distortion risk measures which is defined as distorted expectation of any "non-negative loss"⁴ random variables

$$\rho_g(X) = \int_0^\infty g(1 - F_X(x))dx$$

where the distortion function g is a non-decreasing function with $g(0) = 0$ and $g(1) = 1$ and $g : [0, 1] \rightarrow [0, 1]$, see e.g. [26, 25, 24, 23, 15, 7, 6, 16]. Distortion risk measures were originally applied to a wide variety of insurance problems then later to some asset allocation problems related to the investment risks [14]. The focus is to search a distortion function g so that the resultant distortion risk measure ρ_g satisfies certain properties such as coherence.⁵ In order to generate a coherent risk measure for capital calculation, a *concave* distortion function g is often used to adjust the true probability measure to give more weight to higher risk events. The focus of this method is to construct distortion risk measures ρ_g rather than the distorted distribution itself.

Cobb et al. in [4] introduced multi-modal exponential distributions in the form

$$l(x) = \xi(b) \exp \left[- \int_0^x \frac{g(s)}{v(s)} ds \right] \quad (2.4)$$

where the shape polynomial $g(x)$ has the form $g(x) = b_0 + b_1x + b_2x^2 + \dots + b_kx^k$ and $v(x)$ can be polynomials of degree at most 2, and ξ is a normalisation function. The maximum number of possible modes of l in (2.4) is determined by the degree $k > 0$ of the shape polynomial g . In the special case when $g(x) = b_0 + b_1x + b_2x^2 + b_3x^3$ and $v(x) = 1$, $\lambda = -\frac{b_2}{3b_3}$ is a location parameter, $\sigma = b_3^{-0.25}$ is a scale parameter, $\alpha = \sigma g(\lambda)$ is a measure of skewness and $\beta = -\sigma^2(b_1 + b_2\lambda)$ is a measure of kurtosis. If $(\frac{\alpha}{2})^2 - (\frac{\beta}{3})^3 < 0$, then $l(x)$ in (2.4) is a bimodal exponential density function, and α indicates the relative heights of two modes and β indicates the relative separation of the two modes. In order to generate a bimodal distribution, 4 parameters (i.e. b_0, b_1, b_2, b_3) are needed. It is worth noting that our bimodal Lila distributions defined in Definition 2.1 are sub-exponential, meaning we capture

⁴As we are working with losses defined on $]0, \infty)$, we removed the sign for clarity. As a consequence, the losses are regarded as positive values.

⁵A risk measure satisfying the four axioms of translation invariance, sub-additivity, positive homogeneity, and monotonicity is called coherent, c.f [2].

fatter tails than these exponential distributions introduced in [4], in other words, sub-exponential distribution tails decrease more slowly than any exponential tail).

2.2 Some properties of the Lila distribution

In this subsection, some properties of Lila distributions are analysed, including quantile and the first moment.

Recall G is a S-shape distortion function with its derivative function g ; F_X is the cumulative distribution function of a positive random variable X . Let the newly constructed function L be the cumulative distribution function of a random variable Y , with a density function l .

Quantile: For a given number $q \in [0, 1]$, the q -quantile y_q of the random variable Y is such that $q = L(y_q)$, that is

$$y_q = L^{-1}(q) = (G \circ F)^{-1}(q) = (G^{-1} \circ F^{-1})(q).$$

Mean: By the construction of l in (2.2), we have

$$\begin{aligned} \mathbb{E}[Y] &= \int_0^\infty l(y)ydy = \int_0^\infty (g \circ F_X)(x)f_X(x)xdx \\ &= \mathbb{E}[g(F_X(X))X] \end{aligned} \quad (2.5)$$

where the equality in (2.5) can be seen as the transformation of the random variable X . To find the transformation function $y(x)$ so that $Y = y(X)$, one needs to solve the equation (2.3) for y .

From (2.5), the following result can be obtained:

Proposition 2.1. *(i) If the mean of the random variable X exists, i.e. $\mathbb{E}[X] < \infty$, then the mean of Lila distributions is bounded and the following inequality holds:*

$$\min_{p \in [0,1]} g(p)\mathbb{E}[X] \leq \mathbb{E}[Y] \leq \max_{p \in [0,1]} g(p)\mathbb{E}[X].$$

(ii) Conversely, if the mean of the random variable X does not exist, then the mean of the Lila distribution does not exist.

Proof. (i) To obtain the upper bound of the mean of Lila distribution:

$$\begin{aligned}
\mathbb{E}[Y] &= \int_0^\infty l(y)ydy \\
&= \int_0^\infty x (g \circ F)(x)f(x)dx \\
&\leq \max_{x \in [0, \infty)} (g \circ F)(x) \int_0^\infty xf(x)dx \\
&= \max_{p \in [0, 1]} g(p)\mathbb{E}[X].
\end{aligned}$$

The similar calculation leads to the lower bound:

$$\begin{aligned}
\mathbb{E}[Y] &= \int_0^\infty l(y)ydy \\
&= \int_0^\infty x (g \circ F)(x)f(x)dx \\
&\geq \min_{x \in [0, \infty)} (g \circ F)(x) \int_0^\infty xf(x)dx \\
&= \min_{p \in [0, 1]} g(p)\mathbb{E}[X].
\end{aligned}$$

(ii) The statement holds true due to the way of the construction of Lila distribution. Define $\bar{k} = \arg_x \min g(F_X(x))$. If the mean of the random variable of X does not exist, it means that the tail of X on $[k, \infty)$ is "fat" such that

$$\int_k^\infty f_X(x)dx = \infty$$

for $k > \bar{k}$. Then after the distortion of $f_X(x)$, see Eq. (2.2), the tail of the new random variable Y is even "fatter" and

$$\int_k^\infty l(x)dx = \int_k^\infty g(F_X(x))f_X(x)dx = \infty$$

due to the boundedness of $g(F_X)$. This means the mean of Y is infinite. \square

The above proposition concludes that Lila distributions do not change the infinite mean property of the underlying distributions. In case the mean of the underlying distribution exists, the mean value of the Lila distribution has bounded mean. The bounds of the mean depend on the distortion function G , more precisely on its derivative function g .

3 Examples of the Lila distribution

In this section we demonstrate the construction of the Lila distributions and analyse its basic features with specific underlying distributions F and distortion functions G .

3.1 The main example

Let F_X be the lognormal cumulative distribution function with the location parameter μ and the scale parameter σ . The distortion function G is defined as

$$G(p) = \begin{cases} b - b \left(1 - \frac{p}{a}\right)^{2n+1}, & p \leq a \\ b + (1 - b) \left(\frac{p-a}{1-a}\right)^{2n+1}, & p > a \end{cases} \quad (3.1)$$

with parameters $a \in (0, 1)$, $b \in [0, 1]$ and integer $n \geq 1$. The function G defined in (3.1) is usually referred to as *Double-Odd-Polynomial Seat*. In the case when $n = 1$, (3.1) is referred to as the Double-Cubic Seat ⁶.

The *S*-shape feature of the function G defined in (3.1) are illustrated in Figure 1, Figure 2 and Figure 3. Specifically, set $n = 1$, Figure 1 illustrates the change in the function G by increasing the value of the parameter b from 0 to 1 when the value of the parameter a is fixed. For each graph in Figure 1, we fix the value of the parameter a then we take 7 values of the parameter b . Namely, in each figure, one sees 7 coloured lines corresponding to the 7 values of b .

Similarly, set $n = 1$, Figure 2 illustrates the change in function G by increasing the value of the parameter a when the value of the parameter b is fixed. Figure 3 illustrates G with varying n when $a = 0.5$ and $b = 0.5$.

⁶http://www.flong.com/texts/code/shapers_poly/

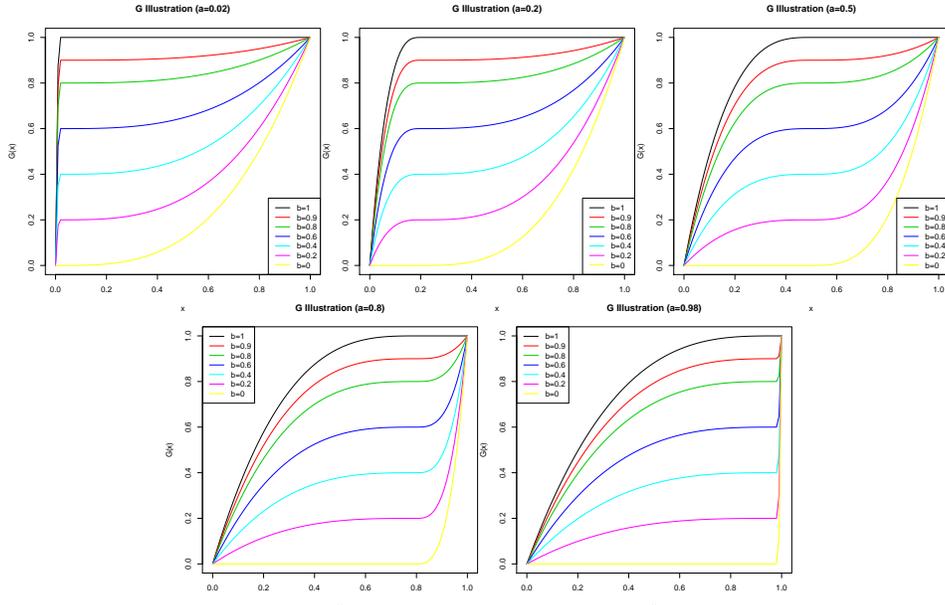


Figure 1: Illustration of G for the fixed parameter a when $n = 1$

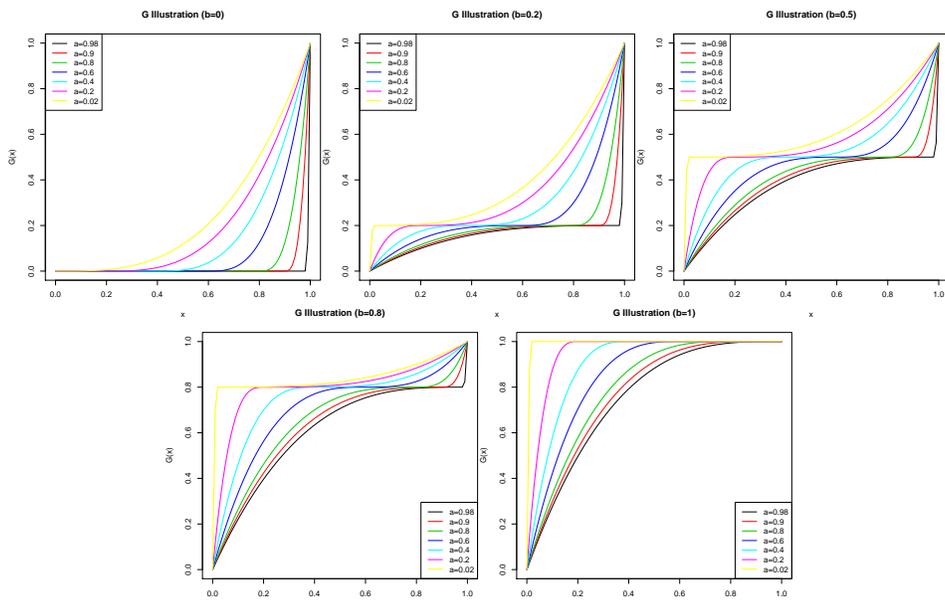


Figure 2: Illustration of G for the fixed parameter b when $n = 1$

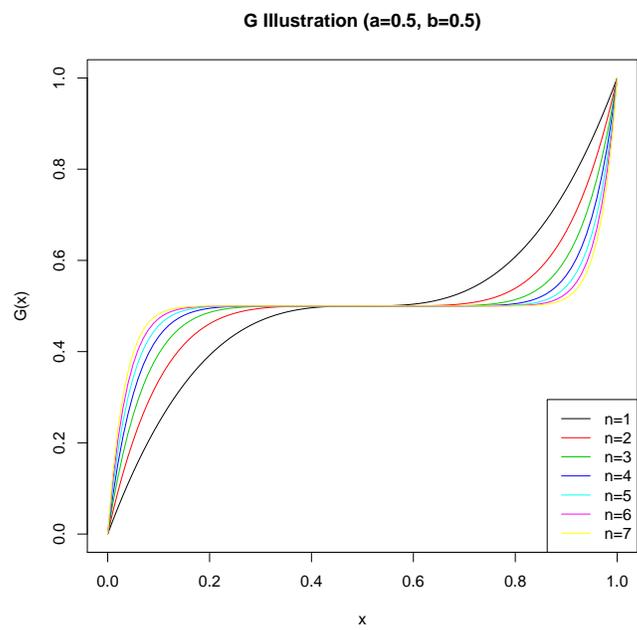


Figure 3: Illustration of G when $a = 0.5$ and $b = 0.5$

In each graph of Figure 1, the parameter a is fixed. Increasing the value of the parameter b moves the function G from the yellow line at the bottom to the black line at the top, namely increasing the height of function G . In each graph of Figure 2, the parameter b is fixed. Increasing the value of the parameter a moves the function G from the yellow line at the left to the black line at the right, meanwhile maintaining the position of the turning point. In Figure 3, increasing the parameter n widens the "flat region".

More insights on the roles of the parameters a , b and n can be obtained by looking at the derivative function g of G :

$$g(p) = \begin{cases} \frac{(2n+1)b}{a^{2n+1}}(p-a)^{2n}, & p \leq a \\ \frac{2n+1(1-b)}{(1-a)^{2n+1}}(p-a)^{2n}, & p > a \end{cases} \quad (3.2)$$

The graphs of the derivative function g are illustrated in Figure 4, Figure 5 and Figure 6. In each graph of Figure 4, the parameter a is fixed and in each graph of Figure 5, the parameter b is fixed, where $n = 1$. One can see that when a is fixed, increasing b lowers the values of the function g on the left and increases the values of the function g on the right; when b is fixed, increasing a shifting the saddle point or turning point of the function g to the right. In Figure 6, one can see that for fixed a and b , increasing n widens the flat region.

Recall that the function g will be used to re-attribute the weight to the original density f and to generate Lila distributions defined in (2.2). The parameter a decides the location where a minimum weight is assigned to; the parameter b controls the scale of weights on each side of the minimum value; the parameter n controls the width of the flat region. For these reasons, the parameter a is referred to as the *turning location* parameter and the parameter b is referred to as the *tail weight* parameter and the parameter n is referred to as the *separate* parameter.

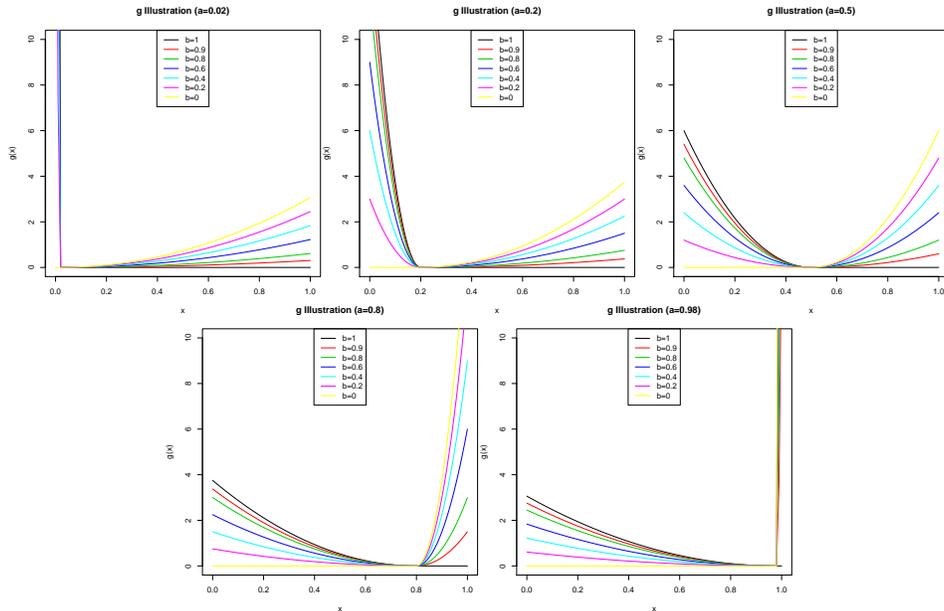


Figure 4: Illustration of g for the fixed parameter a when $n = 1$

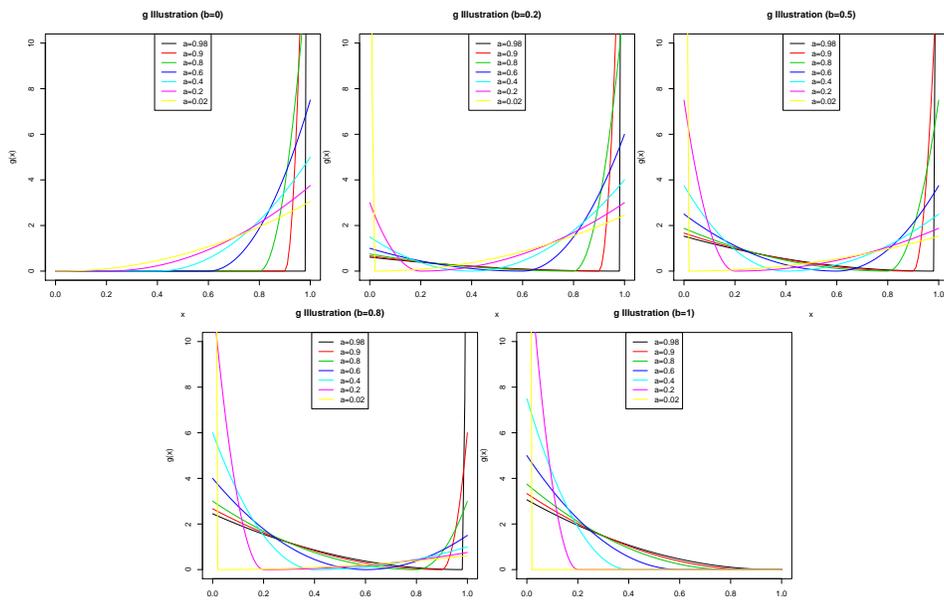


Figure 5: Illustration of g for the fixed parameter b when $n = 1$

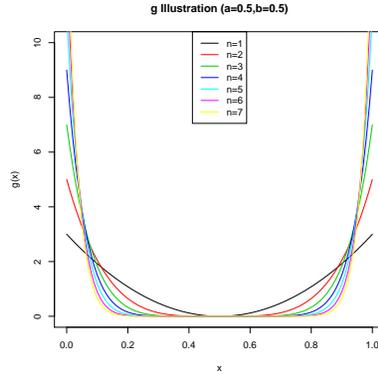


Figure 6: Illustration of g when $a = 0.5$ and $b = 0.5$

Now we are ready to illustrate the Lila distribution L defined in (2.1) and l defined in (2.2). Several graphs of l and L are plotted with different values of the parameters a and b , illustrated in Figure 7, where $\mu = 2$ and $\sigma = 0.5$.

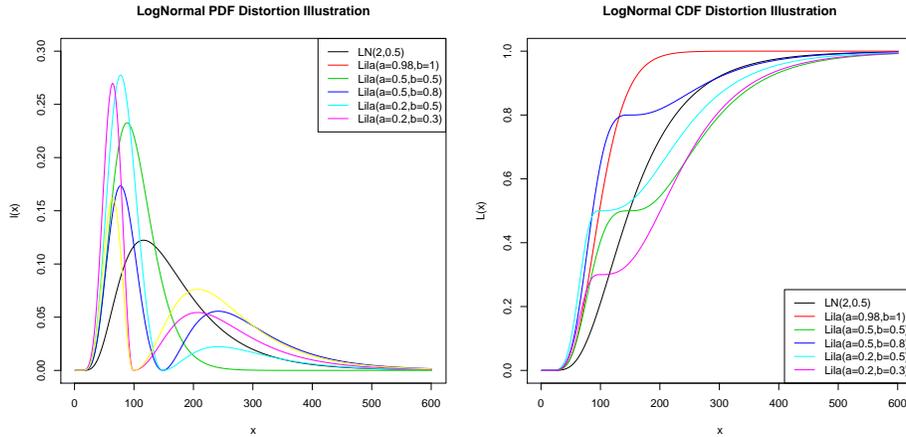


Figure 7: An example of the Lila distribution ($n = 1$)

The example of the probability density function l has a distinctive feature, see in Figure 7: it has two local maximum points, which we refer to as two peaks, the *body peak* on the left and the *tail peak* on the right; the terminology *pseudo bi-modal* comes from this feature: two peaks with different heights.

Next, in the example of G defined in (3.1) and F_X being the lognormal

cumulative distribution function, we investigate what roles parameters a , b and n in function G play in the resulting l .

In Figure 8, the parameters a is fixed in each figure, $\mu = 2$, $\sigma = 0.5$ and $n = 1$, then b takes different values. The various coloured lines on the graph represent each b value.

In Figure 9, the parameters b is fixed in each figure, $\mu = 2$, $\sigma = 0.5$ and $n = 1$, then a takes different values. The various coloured lines on the graph represent each a value.

In Figure 10, $a = 0.5$ and $b = 0.5$, then n takes different values.

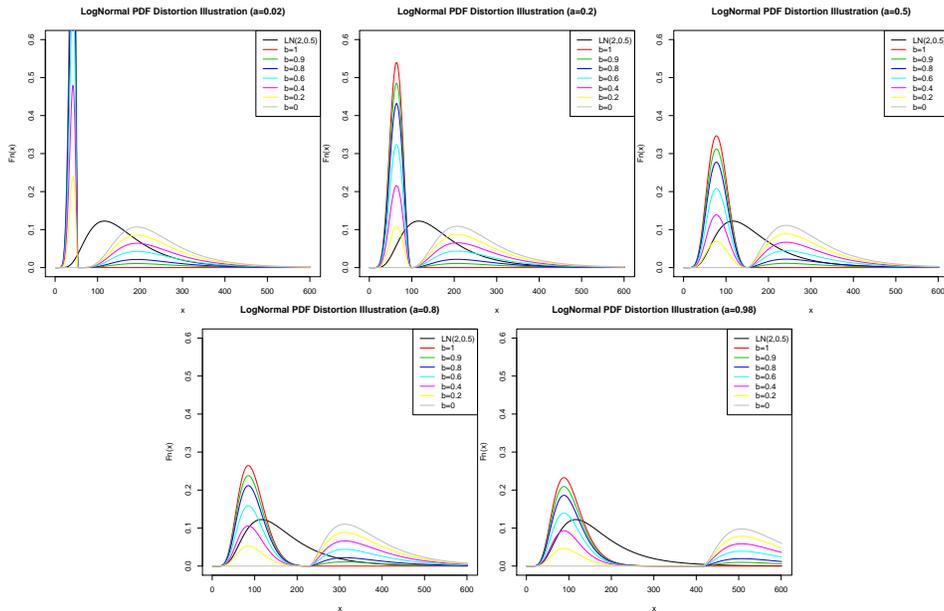


Figure 8: Effect of b varying given a is fixed and $n = 1$

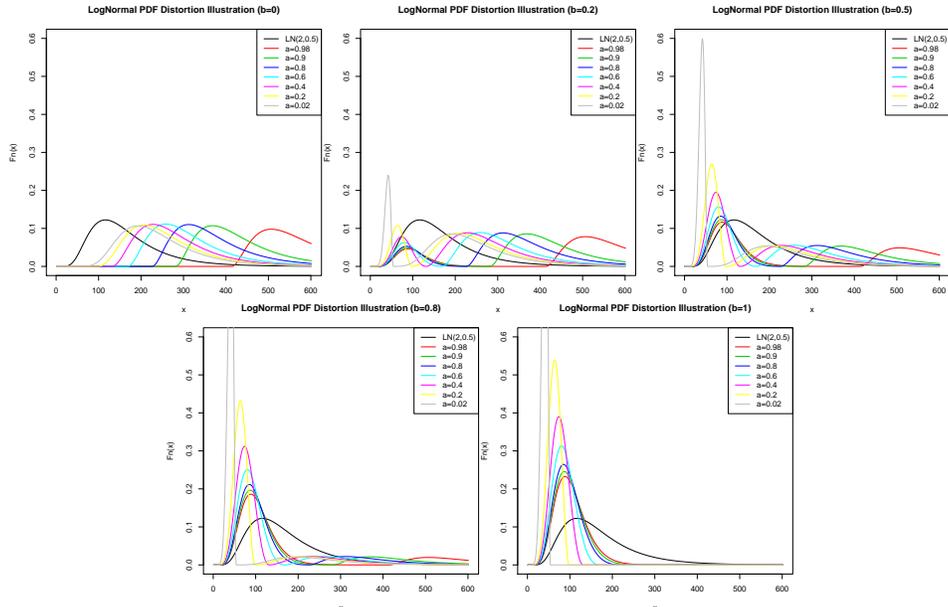


Figure 9: Effect of a varying given b is fixed and $n = 1$

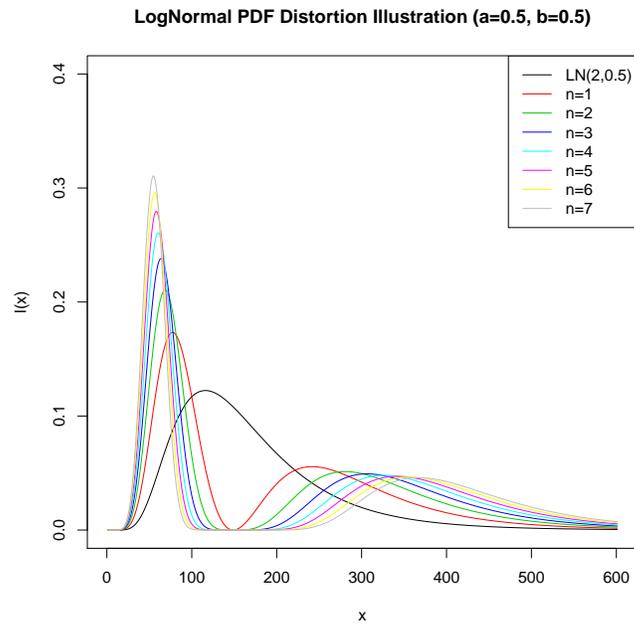


Figure 10: Effect of n varying when $a = 0.5$ and $b = 0.5$

In each of the five graphs in Figure 8, as b increases, g moves from the grey line ($b = 0$), to the yellow line ($b = 0.2$), then to the pink line ($b = 0.4$), then to the light blue line ($b = 0.8$), then to the green line ($b = 0.8$) and finally to the red line ($b = 1$). We observe that the parameter b adjusts the height of the two peaks. As b increases, the height of body peak increases and the height of tail peak decreases. Meanwhile, across the five graphs from left top to bottom right, the parameter a increases from 0.02 to 0.98. The parameter a controls the location and the range of rare events between two peaks.

In each of these five graphs in Figure 9, it is possible to see that the parameter a adjusts the location of and the gaps between the two peaks. As a increases, the range of rare events between two peaks is enlarged and its location shifts to the right; further the height of the tail peak keeps the same for the same s as a change. As a consequence of the preservation of the total mass, the height of the body peak decreases as a increases. Meanwhile, across the five graphs from left top to bottom right, the parameter b increases from 0 to 1. We see that the height of tails decrease as b increases.

In Figure 10, as n increases, the body and tail parts get further apart.

In summary, a Lila distribution constructed with the distortion function G in (3.1) has two peaks of different heights which are call a body peak and a tail peak. The parameters n and a control the range and the location of rare events between two peaks: as a increase the range of rare events between two peaks is enlarged and its location shifts to the right; the parameter b controls the height of tail peaks: as b increases the tail height decreases.

3.2 More examples

One can also consider F_X to be either a Normal distribution function (exponential family) or Weibull distribution (sub-exponential family as the shape parameter is inferior to 1). Figure 11 and 12 exhibit an identical feature (two peaks) as the example above.

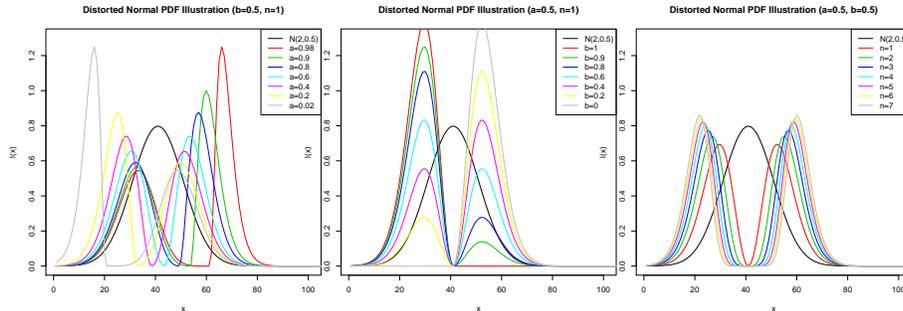


Figure 11: Distorted Normal distribution with G in (3.1)

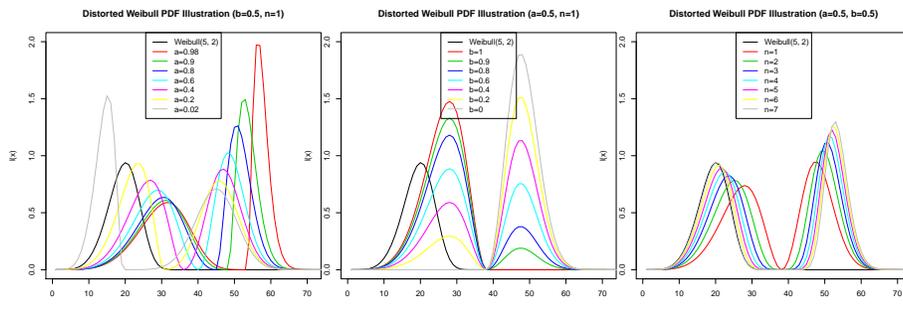


Figure 12: Distorted Weibull distribution with G in (3.1)

Comparing the Gaussian distorted distribution and the Weibull distorted distribution, it is interesting to note that some of the figures look really similar, however the Weibull and the Gaussian distributions are not defined on the same support. Indeed, the Weibull distribution does not take negative values, as a consequence applied to credit and operational risk problems, where the losses are all defined on $]0, \infty)$ (recall the definition of non-negative loss), the use of the Gaussian distribution even distorted would not be appropriate, and the risk measures would mechanically be impacted.

Next, we apply the method introduced in Subsection 2.1 with the function G

$$G(p) = \begin{cases} bp + (1-b)a \left(1 - \left(1 - \frac{p}{a}\right)^{2n+1}\right), & p \leq a \\ bp + (1-b) \left(a + (1-a) \left(\frac{p-a}{1-a}\right)^{2n+1}\right), & p > a \end{cases} \quad (3.3)$$

to construct a Lila distribution with a "shoulder" feature, which is illustrated in Figure 13. In the case when $n = 1$, the G in (3.3) is the Double-Cubic seat with linear blend⁷. Due to a structure similar to the one in (3.1), we will not discuss G in (3.3) any further, and we will only briefly discuss the resulting Lila distributions.

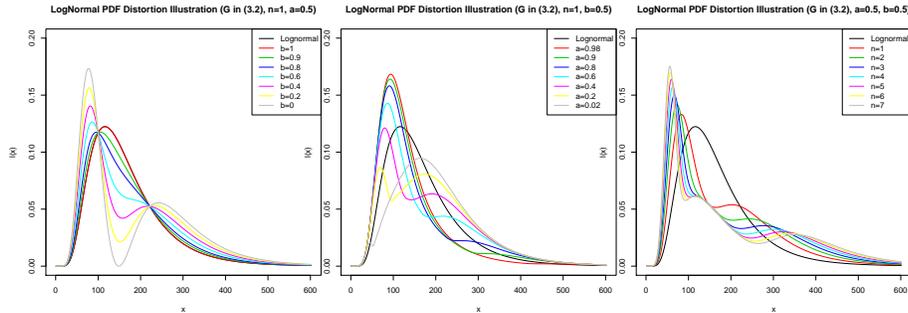


Figure 13: Lila distributions with G in (3.2)

In Figure 13 (a), the parameter a is set to be 0.5. In the case a is fixed, the crossing two points of Lila with lognormal are fixed, say x_1^* and x_2^* with $x_1^* < x_2^*$. In the range (x_1^*, x_2^*) , the density of Lila is lower than lognormal and outside of the range (x_1^*, x_2^*) , the density of Lila is higher than the lognormal. As a decreases from 1 to 0, the gap between the Lila and the lognormal in the region (x_1^*, x_2^*) increases and a "valley" is formed.

⁷http://www.flong.com/texts/code/shapers_poly/

In Figure 13 (b), the parameter b is set to be 0.5. In the case b is fixed, changing a to adjust the shoulder's height and position. As a decreases from 1 to 0, Lila starts to have a shoulder on the right. At certain critical value of a , say $a^* \in (0, 1]$, the two humps have the same height. (The value of a^* depends on the value of b). When a decreases further from a^* to 0, a shoulder on the left start to appear, the main body shifts to the right.

In Figure 13 (c), the parameters $a = 0.5$ and $b = 0.5$. As n increases, the "shoulder" gets lower and two "shoulders" are formed.

It is worth stressing, in order to generate the desired distribution, that we require the derivative function g of G to be continuous, positive and convex (see in subsection 2.1). Here is an example of the function G which does not satisfy this condition:

$$G(p) = \begin{cases} \frac{b}{a} \sqrt{a^2 - (p - a)^2}, & p \leq a \\ 1 - \frac{1-b}{1-a} \sqrt{(1-a)^2 - (p - a)^2}, & p > a. \end{cases} \quad (3.4)$$

This function is often referred to as a Double-Elliptic Seat. Although this example of G produces two "peaks", yet due to the indifferentiability of $g(p)$ at the point $p = a$, the resulting density has some undesired phenomenon of unsmoothness (sudden changes) at certain points.

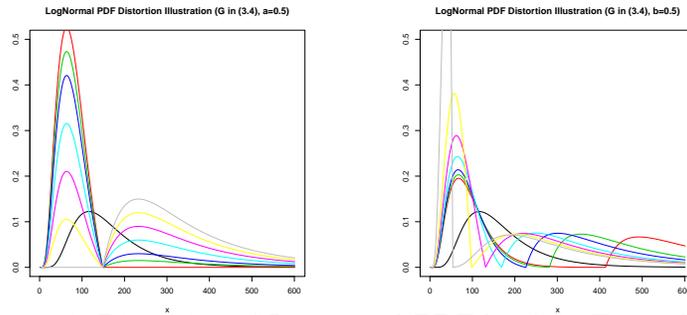


Figure 14: Distortion of Lognormal PDF by G in Example 3.1

4 An application of the Lila distribution

In this section, as an application of the newly introduced Lila distribution, we will fit it to two sets of data. In order to find the best fit parameters of a Lila distribution to the sample data, we apply the quantile distance minimisation method, that is to find the parameters which give the minimum value of the square root of the sum of quantile differences between the empirical distribution and a Lila distribution⁸.

We generate two sets of data: one set of data distinctively has a body peak and a tail peak, see the histogram in Figure 15, and another set of data with a shoulder feature (see the histogram in Figure 16).

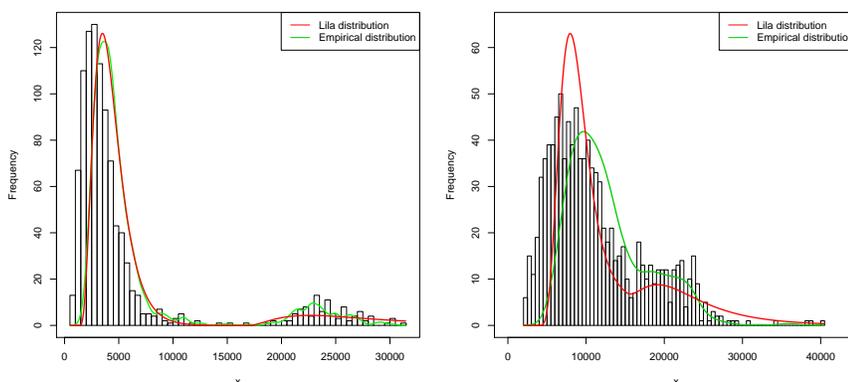


Figure 15: data with 2 local peaks Figure 16: data with 1 shoulder

To fit a Lila distribution, for the data described in Figure 15, the distortion function G in (3.1) with $n = 1$ is used; for the data described in Figure 16, the distortion function G in (3.3) with $n = 1$ is used. In both graphs, the green line represents empirical distributions and the red line the fitted Lila distributions.

5 Conclusion

In this paper, we focus on the presentation of the Lila distribution by combining an S-shape distortion function with a lognormal distribution. The

⁸The maximum likelihood estimation approach has also been tested and is viable in most cases.

newly constructed Lila family of distributions allows capturing multiple peaks or several modes empirical risk data usually have. Yet the Lila family of distributions is neither limited to this kind of distortion function nor to this kind of underlying distributions. Indeed we compared the results with the application of the distortion function to Weibull and Gaussian distributions. The Lila family of distributions could naturally be extended to other distributions and polynomials of order superior to 3 as soon as the resulting composition is a distribution.

References

- [1] AL-Hussaini E. (2012). Composition of cumulative distribution functions. *Journal of Statistical Theory and Applications*, 11(4), 323-336.
- [2] Artzner P., Delbaen F., Eber J.-M. and Heath D. (1999). Coherent Measures of Risk. *Mathematical Finance*, 9(3), 203-228.
- [3] Barndorff-Nielsen O. (1977). Exponentially decreasing distributions for the logarithm of particle size. *Proceedings of the Royal Society of London. Series A, Mathematical and Physical Sciences (The Royal Society)*, 353 (1674), 401-409.
- [4] Cobb L. Koppstein P. and Chen N.H. (1983). Estimation and moment recursion relations for multimodal distributions of the exponential family. *Journal of the American Statistical Association*, 78(381). 124-130.
- [5] Cohen A. C. (1967). Estimation in Mixtures of Two Normal Distributions. *Technometrics*, 9(1), 15-28.
- [6] Denneberg D. (1990). Premium calculation: why standard deviation should be replaced by absolute deviation. *ASTIN Bulletin*, 20(2), 181-90.
- [7] Denneberg D. (1994). Non-additive measure and integral. *Theory and decision library, series B*, Vol. 27. Netherlands: Kluwer Academic Publishers.
- [8] Embrechts, P., Klüppelberg, C., and Mikosch, T. (1997). Modelling Extremal events for insurance and finance. Springer, Berlin.
- [9] Frühwirth-Schnatter, S. (2006). Finite Mixture and Markov Switching Models. Springer, Heidelberg.
- [10] Glyn A.H. (2003). Value-at-Risk: Theory and Practice, San Diego: Academic Press.
- [11] Guégan D. and Hassani B. (2009). A modified Panjer algorithm for operational risk capital calculations. *The Journal of Operational Risk*, 4, 53-72.
- [12] Guégan D. and Hassani B. (2015). Distortion risk measure or the transformation of unimodal distributions into multimodal functions, *Future Perspectives in Risk Models and Finance*, International Series in Operations Research &

Management Science 211, Springer International publishing Switzerland, 77-88.

- [13] Gumbel E.J. (1958). *Statistics of Extremes*, Columbia University Press.
- [14] Hamada M., Sherris M. and van Der Hoek J. (2006). Dynamic portfolio allocation, the dual theory of choice and probability distortion function. *Astin Bulletin*, 36(1), 187-217.
- [15] Hardy, M.R., and Wirth J (2001). Distortion risk measure: coherence and stochastic dominance. working paper, University of Navarra.
- [16] Harlimann W. (2004). Distortion risk measures and economic capital. *North American Actuarial Journal*, 8(1), 86-95.
- [17] Jones M.C. (2004). Families of distributions arising from distributions of order statistics. *Test*, 13(1), 1-43.
- [18] Mao, C.X. (2007). Estimating population sizes for capture-recapture sampling with binomial mixtures. *Computational Statistics & Data Analysis*, 51 (2007) 5211-5219.
- [19] Li J. Q. and Barron A. R. (1999). Mixture Density Estimation, in *In Advances in Neural Information Processing Systems* 12, 279-285.
- [20] Lindsay, B.G. (1995). Mixture models: theory geometry and applications. NSF-CBMS Regional Conference Series in Probability and Statistics, 5. Institute of Mathematical Statistics, Hayward.
- [21] Neslehova, J., Embrechts, P. and Chavez-Demoulin, V. (2006). Infinite Mean Models and the LDA for Operational Risk. *Journal of Operational Risk*, 1(1), 3-25.
- [22] Rockafellar R.T., Uryasev S. (2002). Conditional value-at-risk for general loss distributions. *Journal of Banking & Finance*, 26, 1443-1471.
- [23] Sereda, E. N., Bronshtein, E. M., Rachev, S. T., Fabozzi, F. J., Sun, W., Stoyanov, S. V. (2010). Distortion Risk Measures in Portfolio Optimization, in *Handbook of Portfolio Construction, Volume 3 of Business and Economics*, 649-673.
- [24] Wang S.S.(1995).Insurance pricing and increased limits ratemaking by proportional hazards transforms. *Insurance: Mathematics & Economics*, 17, 43-54.
- [25] Wang S.S. (1996). Premium calculation by transforming the layer premium density. *ASTIN Bulletin*, 26, 71-92.
- [26] Wang S.S. (2000). A class of distortion operators for pricing financial and insurance risks. *Journal of Risk and Insurance*, 67(1), 15-36.