An Advanced Weighted Levy Distribution: Statistical Properties and Application

*Fathima Bi and **Afaq Ahmad

*Research Scholar, Department of Mathematics, Baghwant University, Ajmer, India

**Assistant Professor, Department of Mathematics, Baghwant University, Ajmer, India

Abstract: In order to model price variations in market, finance engineers may employ the concept of Levy distribution. The slow fall off of the Levy distribution model is a good match after price changes. In this paper, a new weighted model is introduced which would be obtained by assigning weights to Levy distribution. This work provides an insight to some basic distributional properties of this distributions such as Moments, moment generating function, Skewness, kurtosis, Shannon's entropy etc. Maximum likelihood estimation and method of moments are employed to estimate the model parameters. For the purpose of illustration the proposed model would be applied to the real data set.

Keywords: Levy distribution, weighted distribution, Maximum likelihood estimation and Shannon's entropy.

1. Introduction

Levy distribution is a continuous probability distribution. In spectroscopy, Levy distribution with frequency as a non-negative and dependent variable is also known as Van-der Waals profile. It is a special case of the inverse gamma distribution.

The density function (pdf) of the Levy distribution is given by

$$f(x,a,c) = \sqrt{\frac{c}{2\pi}}\ \frac{e^{-c/2(x-a)}}{(x-a)^{3/2}} \qquad (1)$$

where a is the location parameter and c is the scale parameter.

Its mean and variance are given by

$$\mu_1' = \mu_2 = \infty \qquad (2)$$

By putting $= 0$, the pdf of one parametric Levy distribution is obtained which is given as

$$f(x,c) = \sqrt{\frac{c}{2\pi}}\ \frac{e^{-c/2x}}{(x)^{3/2}} \qquad (3)$$

2. Weighted Levy distribution

Weighted distribution theory gives an integrated method to study with model design and data interpretation problems. Weighted distributions arise commonly in studies connected to reliability, survival analysis, analysis of family data, biomedicine, ecology and several other areas, see Stene (1981) and Oluyede and George (2002). Several authors would have been presented important consequences on weighted distributions, Rao (1965) had presented a unified model of weighted distribution and known several sampling situations that can show by weighted distributions. These situations occur when the recorded observations cannot be considered as a random sample from the original distributions. This imply in some cases it is not likely to work with a random sample from population. Zelen (1974) presented weighted distribution to represent what is called as a length-biased sampling. Patil and Ord (1976) studied a size biased sampling and related invariant weighted distributions. Gupta and Tripathi (1996) studied the weighted version of the bivariate logarithmic series distribution, which has applications in many fields such as: ecology, social and behavioural sciences. Ahmed et al. (2016) discussed length biased weighted Lomax distribution with its applications.

To existent the idea of a weighted distribution, suppose that X is a nonnegative random variable with its probability density function (pdf) $f(x)$, then the p.d.f. of the weight random variable X_w is known by

$$f_w(x) = \frac{w(x)f(x)}{E\left(w(x)\right)} \qquad x \geq 0 \qquad (4)$$

where $w(x)$ be a non-negative weight function.

Depending upon the choice of the weight function $w(x)$, we have different weighted models.

The one parametric weighted Levy distribution is obtained by taking the weights x^{Θ}, to the one parametric levy distribution.

In this paper, the one parametric weighted Levy distribution is proposed with pdf

$$f(x, \Theta, c) = \frac{x^{\theta-3/2}\left(\frac{c}{2}\right)^{\frac{1}{2}-\theta} e^{-c/_{2x}}}{\Gamma(-\theta+1/2)} \qquad (5)$$

where Θ and c are scale parameter and $\int_0^{\infty} f(x, \Theta, c)dx = 1$

3. Statistical Properties

In this section we shall discuss structural properties of one parametric weighted Levy distribution, especially mode, moments, coefficient of variation, moment generating function, skewness, and kurtosis.

3.1 Moments

Suppose X denote the random variable of one parametric weighted Levy distribution with parameters Θ and c, then

$$E(X^r) = \mu_r' = \int_0^\infty x^r f(x,\Theta,c)dx$$

$$= \frac{(c/2)^{\frac{1}{2}-\Theta}}{\Gamma(-\theta+1/2)} \int_0^\infty x^{r+\Theta-3/2} e^{-c/_{2x}} dx$$

$$= \frac{(c/2)^r}{\Gamma(-\theta+1/2)} \int_0^\infty t^{-r-\Theta-1/2} e^{-t} dx$$

$$\mu_r' = \frac{(c/2)^r \Gamma(-r-\theta+1/2)}{\Gamma(-\theta+1/2)} \qquad (6)$$

Substituting r = 1,2,3,4 we get first four moments

Mean = $\mu_1' = \frac{-c}{2\theta+1}$

$$\mu_2' = \frac{c^2}{(2\theta+1)(2\theta+3)}$$

$$\mu_3' = \frac{-c^3}{(2\theta+1)(2\theta+3)(2\theta+5)}$$

Variance $= \mu_2 = \mu_2' - (\mu_1')^2 = \frac{-2c^2}{(2\theta+1)^2(2\theta+3)}$

Standard Deviation $\boldsymbol{\sigma} = \frac{\boldsymbol{ic\sqrt{2}}}{(2\theta+1)\sqrt{(2\theta+3)}}$

Coefficient of Variation $C.V = \frac{-i}{\sqrt{(\theta+3/2)}}$

3.2 Moment generating function

In this sub section we derived the moment generating function of one parametric weighted Levy distribution. From the definition of moment generating function we have

$$M_x(t) = E(e^{tx}) = \int_0^\infty e^{tx} f(x,\Theta,c)dx$$

$$= \int_0^\infty \sum_{j=0}^\infty \frac{(tx)^j}{j!} f(x,\Theta,c)dx$$

$$= \sum_{j=0}^\infty \frac{(t)^j}{j!} \int_0^\infty x^j f(x,\Theta,c)dx$$

$$=\sum_{j=0}^{\infty}\frac{(t)^j}{j!}\mu_j'$$

$$\Longrightarrow M_X(t)=\sum_{j=0}^{\infty}\frac{(t)^j}{j!}\frac{(c/2)^j\Gamma(-j-\theta+1/2)}{\Gamma(-\theta+1/2)} \tag{7}$$

3.3 Skewness and Kurtosis

Coefficient of skewness one parametric weighted Levy distribution is given by

$$\boldsymbol{Sk}=\frac{\mathbf{2}}{\boldsymbol{c}^{\mathbf{3}}}\left(\frac{(\mathbf{2}\boldsymbol{\theta}+\mathbf{1})^{\mathbf{3}}}{\mathbf{2}\boldsymbol{\theta}+\mathbf{5}}\right) \tag{8}$$

Kurtosis of one parametric weighted Levy distribution is given by

$$KR=\frac{-c^2(8\theta^3+16\theta^2-152\theta-382)}{(2\theta+1)(2\theta+5)(2\theta+7)} \tag{9}$$

3.4 Mode

In order to discuss monotonicity of one parametric WLD, we take the logarithm of its pdf as follows:

$$\log\big(f(x,\Theta,c)\big)=(\theta-3/2)\log x+\left(\frac{1}{2}-\theta\right)log\ c/2-\frac{c}{2x}-log\Gamma(-\theta+c/2)$$

Differentiating the above equation with respect to *x* and equating to zero, we obtain

$$x=\frac{c}{3-2\theta} \tag{10}$$

4. Estimation of parameter

In this section, we derive the estimates of parameters of weighted Levy distribution by various methods of estimation viz method of moments and maximum likelihood estimation.

4.1 Methods of Moments

Replacing sample moment with population moments, we get

$$\frac{1}{n}\sum_{i=1}^{n}x_i=\mu_1'$$

$$\Longrightarrow\ \bar{x}=\frac{-c}{2\theta+1}$$

$$\Longrightarrow c=-\bar{x}\,(2\theta+1) \tag{11}$$

and $\frac{1}{n}\sum_{i=1}^{n}{x_i}^2=\mu_2'$

$$\Longrightarrow\ \frac{1}{n}\sum_{i=1}^{n}{x_i}^2-\bar{x}^2=\mu^2$$

$$\hat{\theta}=\frac{3/n\sum_{i=1}^{n}{x_i}^2-\bar{x}^2}{2\left(\bar{x}^2-\sum_{i=1}^{n}{x_i}^2\right)} \tag{12}$$

Substituting the value of $\hat{\theta}$ in (10) we get estimate of c which is given by,

$$\Longrightarrow \hat{c} = \frac{\left(\frac{3}{n}-1\right)\sum_{i=1}^{n} x_i{}^2-\bar{x}^2}{\sum_{i=1}^{n} x_i{}^2-\bar{x}^2} \tag{13}$$

4.2 Method of Maximum Likelihood Estimator

The method Maximum likelihood estimation is the most popular technique used for estimating the parameters of one parametric Levy distribution. Let $x_1, x_2, x_3, \ldots\ldots. x_i$ be a random sample from the one parametric Weighted Levy distribution, then the corresponding log likelihood function is given by,

$$l(\theta) = log\ L(X, \theta, c) = \sum_{i=0}^{n-1} logf(x_i, \theta, c)$$

$$=\sum_{i=0}^{n-1} log\left(\frac{x^{\theta-3/2}\left(\frac{c}{2}\right)^{\frac{1}{2}-\theta} e^{-c/_{2x}}}{\Gamma(-\theta+1/2)}\right)$$

$$log\ L(X, \theta, c) = \left(\frac{1}{2}-\theta\right) log(c/2)n + (\theta-3/2)\sum_{i=0}^{n-1} log\,(x_i) + c/2\sum_{i=0}^{n-1}\frac{1}{x_i} - nlog\Gamma(-\theta+1/2) \tag{14}$$

Now differentiating above with respect to the parameters, we obtain the normal equations

$$-nlog\frac{c}{2} + \sum_{i=0}^{n-1} logx_i + n\,\frac{\Gamma'(-\theta+1/2)}{\Gamma(-\theta+1/2)} = 0$$

$$\Longrightarrow\ \varphi(\theta) = log\frac{c}{2} - 1/n\ \sum_{i=0}^{n-1} logx_i \tag{15}$$

where $\varphi(\theta) = \frac{\Gamma'(-\theta+1/2)}{\Gamma(-\theta+1/2)}$

and

$$c = \frac{(1-2\theta)}{1/n\ \sum_{i=0}^{n-1} logx_i} \tag{16}$$

Solving equations (15) and (16), we get the MLE's of parameters as given below

$$\hat{\theta} = log\frac{c}{2} - 1/n\ \sum_{i=0}^{n-1} logx_i \tag{17}$$

and

$$\hat{c} = \frac{(1-2\hat{\theta})}{1/n\ \sum_{i=0}^{n-1} logx_n} \tag{18}$$

5. Shannon's Entropy

The Shannon entropy equation provides a way to estimate the average minimum number of bits needed to encode a string of symbols, based on the frequency of the symbols. Shannon entropy provides a lower bound for the compression that can be achieved by the data representation (coding) compression step. Shannon entropy makes no statement about the compression efficiency that can be achieved by predictive compression. Algorithmic complexity (Kolmogorov complexity) theory deals with this area. Given an infinite data set (something that only mathematicians possess), the data set can be examined for randomness. If the data set is not random, then there is some program that will generate or approximate it and the data set can, in theory, be compressed.

Shannon Entropy is *H(x)* for one parametric WLD is given by the formula

$$H(x) = E[-logf(x,\Theta,c)]$$

$$= \int_0^\infty -logf(x,\Theta,c)\, f(x,\Theta,c)dx$$

$$=-\int_0^\infty \left(\theta-\frac{3}{2}\right) logx\, f(x,\Theta,c)\, dx + \left(\frac{1}{2}-\theta\right)\log\frac{c}{2}\int_0^\infty f(x,\Theta,c)\, dx -$$

$$\frac{c}{2}\int_0^\infty x\, f(x,\Theta,c)\, dx - \log\Gamma(-\theta+1/2)\int_0^\infty f(x,\Theta,c)\, dx$$

Therefore we get,

$$H(x) = -\left(\theta-\frac{3}{2}\right) E[logx] - \left(\frac{1}{2}-\theta\right)\log\frac{c}{2} + \frac{c^2}{2(2\theta+1)} + \log\Gamma(-\theta+1/2) \qquad (19)$$

References:

- Ahmad, A., Ahmad, S.P. and Ahmed, A., (2016). “Length Biased Weighted Lomax distribution: Statistical Properties and Applications”, *Pakistan Journal of Statistics and Operations Research*, vol. 12(2), pp. 245-255.

- Fisher, R.A., 1934. “The effect of methods of ascertainment upon the estimation of frequencies”, Ann. Eng. 6, 13–25.

- Patil, G.P., Rao, C.R., (1978). Weighted distributions and size-biased sampling with applications to wildlife and human families. Biometrics 34, 179–189.

- Patil, G.P., Rao, C.R., Ratnaparkhi, M.V., (1986). “On discrete weighted distributions and their use in model choice for observed data”, Communication in Statistics - Theory and Methods, vol. 15(3), 907–918.

- Rao, C.R., (1965). “On discrete distributions arising out of methods of ascertainment. Classical and Contagious Discrete Distributions”, Eds. G.P. Patil Pergamon Press and Statistical Publishing Society -Calcutta.

- Gupta AK, Tripathi RC (1996) Weighted bivariate logarithmic series distributions. Commun Stat Theory Methods 25:1099–1117.

- Patil GP, Ord JK (1976) On size biased sampling and related form invariant weighted distribution. Indian J Stat 39:48–61.

- Stene J (1981) Probability distributions arising from the ascertainment and the analysis of data on human families and other groups. Stat Distrib Sci Work Appl Phys Soc Life Sci 6:233–244.

- Zelen M (1974) Problems in cell kinetics and the early detection of disease. In: Proschan F, Sering RJ (eds) Reliability and biometry. SIAM, Philadelphia.