

Uniform Minimum Variance Unbiased Estimator of Fractal Dimension

Zeny L. Maureal^{1*}, Elmer C. Castellano², Roberto N. Padua³

¹University of Science and Technology in Southern Philippines, Cagayan de Oro City, Philippines,
<https://orcid.org/0000-0002-3270-0349>

²University of Science and Technology in Southern Philippines, Cagayan de Oro City, Philippines,
<https://orcid.org/0000-0002-1230-8817>

³University of Science and Technology in Southern Philippines, Philippines,
<https://orcid.org/0000-0002-2054-0835>

*Email Correspondence: zmaureal2014@gmail.com

Abstract

The paper introduced the concept of a fractal distribution using a power-law distribution. It proceeds to determining the relationship between fractal and exponential distribution using a logarithmic transformation of a fractal random variable which turns out to be exponentially distributed. It also considered finding the point estimator of fractional dimension and its statistical characteristics. It was shown that the maximum likelihood estimator of the fractional dimension λ is biased. Another estimator was found and shown to be a uniformly minimum variance unbiased estimator (UMVUE) by Lehmann-Scheffe's theorem.

Keywords: fractal dimension, fractal distribution, uniform minimum variance unbiased estimator

1.0 Introduction

Fractals are normally associated with geometry in the tradition of Mandelbrot (1983). Likewise, in an analytic sense Mandelbrot and Van Ness (1968) described fractals as continuous, but nowhere differentiable functions such as the path traced by a standard Brownian process $\{\beta_t; t \geq 0\}$. In this paper, we considered yet another characterization of fractals; namely as data with more smaller values than larger ones (e.g. income distribution). This idea of migrating from fractal geometry to fractal statistics was uncovered by Padua and Borres (2013). A natural framework for studying these data is to assume a Pareto distribution:

$$f(x) = \frac{\lambda-1}{\theta} \left(\frac{x}{\theta}\right)^{-\lambda}, x \geq \theta > 0, \lambda > 1 \quad (1)$$

which we shall henceforth refer to as a fractal distribution with dimension λ , where λ takes values between integral values, and θ is the minimum of random variables $\{x_1, x_2, \dots, x_n\}$. Equation (1) was derived by Padua et al. (2013) using the characteristics of fractals, self-similarity and fractional dimension. This function is self-similar, that is, it retains the same shape at all scales:

$$f(ax) = a^{-\lambda} f(x) \quad (2)$$

hence, the name fractal distribution. Given a random sample $\{x_1, x_2, \dots, x_n\}$ from (1), we are interested in finding the statistical characteristics about the estimator of fractal dimension λ .

2.0 Fundamental Concepts

The following theorems are the properties of power law distribution or fractal distribution (Clauset et al., 2009; Newman, 2005):

Theorem 1. Let X be fractal random variable with probability distribution given in (1). Its moments are only well defined for $m < \lambda - 1$. For $m \geq \lambda - 1$, all moments diverge.

Proof. Let X be distributed as power-law. Its moments are given by

$$\langle x^m \rangle = \int_{\theta}^{\infty} x^m f(x) dx$$

$$\langle x^m \rangle = \frac{\lambda - 1}{\lambda - 1 - m} \theta^m$$

It is well defined for $m < \lambda - 1$. ■

Theorem 2. For $\lambda > 2$, the mean is perfectly well defined and its variance is finite when $\lambda > 3$.

Proof. For power law distribution, the mean is

$$\mu = E(X) = \frac{\lambda - 1}{\theta^{1-\lambda}} \int_{\theta}^{\infty} x^{1-\lambda} dx$$

$$\mu = \frac{\lambda - 1}{\lambda - 2} \cdot \theta. \text{ This is defined when } \lambda > 2.$$

Since $\sigma^2 = E(x^2) - \mu^2$, then $\sigma^2 = \frac{\lambda - 1}{\lambda - 3} \theta^2 - \frac{(\lambda - 1)^2}{(\lambda - 2)^2} \cdot \theta^2$

or $\sigma^2 = \frac{\lambda - 1}{(\lambda - 3)(\lambda - 2)^2} \cdot \theta^2$. It has finite variance when $\lambda > 3$. ■

Most identified power laws in nature have exponents such that the mean exists, but the variance does not.

Theorem 3. For $\lambda > 1$, median always exists.

Proof. For X distributed as power law, the median m is given by

$$\int_{\theta}^m f(x) dx = \frac{1}{2}$$

$$\frac{\lambda - 1}{\theta} \int_{\theta}^m \left(\frac{x}{\theta}\right)^{-\lambda} dx = \frac{1}{2}$$

$$\frac{\lambda - 1}{\theta^{1-\lambda}} \cdot \frac{x^{1-\lambda}}{1-\lambda} \Big|_{\theta}^m = \frac{1}{2}$$

$$\frac{m^{1-\lambda}}{\theta^{1-\lambda}} = \frac{1}{2}$$

$$m = 2^{\frac{1}{\lambda-1}} \cdot \theta$$

For $\lambda > 1$, median always exists and is defined. ■

One way of determining the estimate of the exponent of a power-law distribution is by maximum likelihood estimation (Malik, 1970; Ramachandran & Tsokos, 2009; Bauke, 2007). Padua et al. (2015) derived the maximum likelihood estimator of λ .

Theorem 4. The maximum likelihood estimator of λ is

$$\hat{\lambda} = 1 + \frac{1}{\frac{\sum_{i=1}^n \ln \frac{x_i}{\theta}}{n}}$$

Where $\theta = \min\{x_1, x_2, \dots, x_n\}$

Proof.

Let $x_i \sim F(x, \lambda)$. The likelihood function is given by

$$L = \prod_{i=1}^n \frac{\lambda - 1}{\theta} \left(\frac{x_i}{\theta}\right)^{-\lambda} \tag{3}$$

It is easier to work when we take the natural logarithm of the likelihood function.

$$\ln L = \sum_{i=1}^n \left[\ln(\lambda - 1) - \ln \theta - \lambda \ln \frac{x_i}{\theta} \right]$$

$$\ln L = n \ln(\lambda - 1) - n \ln \theta - \lambda \sum_{i=1}^n \ln \frac{x_i}{\theta}$$

Then, we maximize the log-likelihood by taking the first derivative with respect to λ and equate to zero.

$$\frac{d \ln L}{d \lambda} = \frac{n}{\lambda - 1} - \sum_{i=1}^n \ln \frac{x_i}{\theta} = 0$$

Solving for $\hat{\lambda}$:

$$\hat{\lambda}_{MLE} = 1 + \frac{n}{\sum_{i=1}^n \ln \frac{x_i}{\theta}} \text{ or } \hat{\lambda}_{MLE} = 1 + \frac{1}{\frac{\sum_{i=1}^n \ln \frac{x_i}{\theta}}{n}}$$

3.0 Relationship of Fractal Distribution and Exponential Distribution

The distribution (1) is related to the more familiar exponential distribution as contained in the following theorem:

Theorem 5. Let X have the probability density (1). Let $y = \log \left(\frac{x}{\theta}\right)$, then Y has the exponential distribution:

$$g(y) = \beta e^{-\beta y}, y > 0 \text{ where } \beta = \lambda - 1. \tag{4}$$

Proof. Let $y = \log\left(\frac{x}{\theta}\right)$, then $x = \theta e^y$ and $\frac{dx}{dy} = J = \theta e^y$, is the Jacobian transformation. Thus,

$$g(y) = f\left(\log\left(\frac{x}{\theta}\right)\right) \cdot J$$

$$g(y) = \frac{(\lambda - 1)}{\theta} \left(\frac{\theta e^y}{\theta}\right)^{-\lambda} \cdot \theta e^y$$

$$g(y) = (\lambda - 1)e^{-(\lambda-1)y}$$

$$g(y) = \beta e^{-\beta y}, \quad y \geq 0. \blacksquare$$

Next, if $Y = \log X$ is distributed as exponential with parameter $\beta = \lambda - 1$, then we need to know the distribution of X in the following theorem:

Theorem 6. Let $Y = \log X$. If Y is distributed as exponential with $\beta = \lambda - 1$, then X has fractal distribution given in (1).

Proof. Let $Z = \log X$ and $J = \left|\frac{dz}{dx}\right| = \frac{1}{x}$. Then,

$$f(x) = (\lambda - 1)e^{-(\lambda-1)\log x}.$$

$$\frac{1}{x} = (\lambda - 1)e^{\log x^{-(\lambda-1)}}.$$

$$\frac{1}{x} = (\lambda - 1)x^{1-\lambda}. \quad x^{-1} = (\lambda - 1)x^{-\lambda}$$

Thus, X is fractal with $\theta = 1$. \blacksquare

From Theorem 5 and 6, we established the fundamental theorem of Fractal Statistics.

Theorem 7. Fundamental Theorem for Fractal Statistics

X is fractal if and only if $\log\left(\frac{x}{\theta}\right)$ has an exponential distribution with parameter $\beta = \lambda - 1$

Proof:

Proof follows from Theorem 5 and Theorem 6. \blacksquare

The fundamental theorem clearly connects the power-law distribution with an exponential distribution. The properties of random variables with exponential distribution are already well-known (Feller, 1991; Ross, 2009; Temme, 2010). With this, we can establish the properties of random variables having power law distribution.

If Y is exponential, then we can also obtain the distribution of $\sum_{i=1}^n Y_i$.

Theorem 8. Let y_1, y_2, \dots, y_n be iid $\exp(\beta)$, then

$$Z = \sum_{i=1}^n Y_i \tag{5}$$

distributed as Gamma $\left(n, \frac{1}{\beta}\right)$.

Proof. The moment generating function of Y_i is $m_{y_i}(t) = \frac{1}{1-\beta t}$, and so:

$$m_z(t) = \prod_{i=1}^n \frac{1}{1-\beta t} = \frac{1}{(1-\beta t)^n} \tag{6}$$

which is the moment generating function of Gamma $(\alpha = n, \beta t)$ \blacksquare

4.0 Properties of Estimator of Fractal Dimension

Let y_1, y_2, \dots, y_n be iid $\exp(\beta)$. The likelihood function is

$$L(y_1, y_2, \dots, y_n, \beta) = \beta^n e^{-\beta \sum_{i=1}^n y_i} \tag{7}$$

Then the maximum likelihood estimator of β is given by

$$\hat{\beta} = \frac{1}{\bar{y}} \tag{8}$$

However, $\hat{\beta}$ is a biased estimator β . Regalado et al. (2018) derived the unbiased estimator for β and the estimator is given by

$$\hat{\beta} = \frac{n-1}{n} \cdot \frac{1}{\bar{y}} \tag{9}$$

5.0 Properties of Estimator of Fractal Distribution

Using the estimator of λ in Theorem 4, we find the expected value of $\hat{\lambda}$ as follows:

$$E(\hat{\lambda}) = E\left(1 + \frac{1}{\bar{y}}\right), \text{ where } \bar{y} = \frac{\sum_{i=1}^n \log X_i}{n}$$

Letting $s = \sum_{i=1}^n \log X_i$, we get

$$E(\hat{\lambda}) = 1 + nE(s^{-1})$$

Calculating $E(s^{-1})$, where s is distributed as Gamma (Theorem 8), we have

$$E(s^{-1}) = \int_0^\infty \frac{\beta^n}{\Gamma(n)} s^{n-2} e^{-\beta s} ds$$

$$E(s^{-1}) = \frac{\beta}{n-1} \int_0^\infty \frac{\beta^{n-1}}{\Gamma(n-1)} s^{n-2} e^{-\beta s} ds.$$

The intergral $\int_0^\infty \frac{\beta^{n-1}}{\Gamma(n-1)} s^{n-2} e^{-\beta s} ds$

is a gamma distribution with parameters $\alpha = n - 1$ and $\beta > 0$. Thus,

$$E(s^{-1}) = \frac{\beta}{n-1}$$

By (4),

$$E(s^{-1}) = \frac{\beta}{n-1} = \frac{\lambda-1}{n-1} \tag{10}$$

Thus,

$$E(\hat{\lambda}) = 1 + \frac{n}{n-1}(\lambda - 1) \tag{11}$$

Here, $E(\hat{\lambda}) \neq \lambda$. $\hat{\lambda}$ is a biased estimator of λ . However, as n gets large, $\hat{\lambda}$ is a consistent estimator for λ .

Theorem 9. Let X be distributed as fractal. The unbiased estimator $\hat{\lambda}$ of λ is

$$\hat{\lambda} = 1 + \frac{n-1}{n} \cdot \frac{1}{\bar{y}} \text{ where } \bar{y} = \frac{\sum_{i=1}^n \log X_i}{n}$$

Proof. By Theorem 4, $\log X_i$ is distributed as exponential with $\beta = \lambda - 1$. Thus, by (9)

$$\hat{\lambda} = 1 + \hat{\beta} = 1 + \frac{n-1}{n} \cdot \frac{1}{\bar{y}}$$

Next, we find the bias by finding $E(\hat{\lambda})$.

$$E(\hat{\lambda}) = E\left(1 + \frac{n-1}{n} \cdot \frac{1}{\bar{y}}\right) = 1 + \frac{n-1}{n} E\left(\frac{1}{\bar{y}}\right)$$

Letting $s = \sum_{i=1}^n \log X_i$, we get

$$E(\hat{\lambda}) = 1 + \frac{n-1}{n} \cdot nE(s^{-1})$$

$$\text{By (10), } E(\hat{\lambda}) = 1 + (n-1) \cdot \frac{\lambda-1}{n-1} = \lambda$$

By definition of bias, $Bias = E(\hat{\lambda}) - \lambda = 0$. Hence, $\hat{\lambda}$ is an unbiased estimator of λ . ■

Theorem 10. Let x_1, x_2, \dots, x_n be random variables distributed as fractal with parameter λ . The MLE of λ is consistent estimator for λ .

Proof. The maximum likelihood estimator for λ is given in Theorem 4. Since the MLE is a biased estimator, we will show that

$$\lim_{n \rightarrow \infty} E\left[(\hat{\lambda} - \lambda)^2\right] = 0$$

Since $E\left[(\hat{\lambda} - \lambda)^2\right] = Var(\hat{\lambda}) + Bias^2(\hat{\lambda})$, then we need to show that

$$\lim_{n \rightarrow \infty} [Var(\hat{\lambda}) + Bias^2(\hat{\lambda})] = 0$$

or we show that

$$\lim_{n \rightarrow \infty} Var(\hat{\lambda}) = 0 \text{ and } \lim_{n \rightarrow \infty} Bias(\hat{\lambda}) = 0$$

Find $Var(\hat{\lambda})$:

$$Var(\hat{\lambda}) = Var\left(1 + \frac{1}{\bar{y}}\right) = Var\left(\frac{1}{\bar{y}}\right) = Var\left(\frac{n}{\sum_{i=1}^n y_i}\right) = n^2 Var(s^{-1})$$

$$Var(s^{-1}) = E(s^{-2}) - (E(s^{-1}))^2$$

Find $E(s^{-2})$:

$$E(s^{-2}) = \int_0^\infty \frac{\beta^n}{\Gamma(n)} s^{n-3} e^{-\beta s} ds$$

$$E(s^{-2}) = \frac{\beta^2}{(n-1)(n-2)} \int_0^\infty \frac{\beta^{n-2}}{\Gamma(n-2)} s^{n-3} e^{-\beta s} ds,$$

$$\int_0^\infty \frac{\beta^{n-2}}{\Gamma(n-2)} s^{n-3} e^{-\beta s} ds = 1 \text{ is a Gamma}$$

distribution with parameters $\alpha = n - 1$ and $\beta > 0$

$$E(s^{-2}) = \frac{\beta^2}{(n-1)(n-2)} \tag{12}$$

With $\beta = \lambda - 1$,

$$Var(s^{-1}) = \frac{(\lambda-1)^2}{(n-1)(n-2)} - \frac{(\lambda-1)^2}{(n-1)^2}$$

$$Var(s^{-1}) = \frac{1}{(n-1)^2(n-2)} (\lambda - 1)^2$$

Thus, $Var(\hat{\lambda}) = \frac{n^2}{(n-1)^2(n-2)} (\lambda - 1)^2$, where the

coefficient $\frac{n^2}{(n-1)^2(n-2)} < 1$. As $n \rightarrow \infty$,

$$Var(\hat{\lambda}) \rightarrow 0, \text{ i.e., } \lim_{n \rightarrow \infty} Var(\hat{\lambda}) = 0$$

Next, we find the Bias($\hat{\lambda}$):

$$Bias(\hat{\lambda}) = E(\hat{\lambda}) - \lambda$$

$$Bias(\hat{\lambda}) = 1 + \frac{n}{n-1}(\lambda - 1) - \lambda = \frac{n-1+n\lambda-n-n\lambda+\lambda}{n-1}$$

$$Bias(\hat{\lambda}) = \frac{\lambda-1}{n-1}$$

As $n \rightarrow \infty$, i.e., $\lim_{n \rightarrow \infty} Bias(\hat{\lambda}) = 0$. ■

Theorem 11. $\hat{\lambda}$ has sufficient and complete statistic for λ .

Proof. Using (3), by Neyman-Fisher factorization theorem $T = \sum_{i=1}^n y_i$ is a sufficient statistic for λ .

Next, we showed completeness:

Let $g(t)$ be a function of statistic t such that $E(g(t)) = 0$. Then $E(g(t)) = \int_0^\infty g(t) \cdot \beta e^{-\beta t} dt = 0$.

Since $\beta > 0$, then $\beta e^{-\beta t} \neq 0$. This implies that

$g(t) = 0, \forall t$. Therefore, T is a complete sufficient

statistic. ■

Since $\hat{\lambda}$ a function of complete sufficient statistic, then we can say that it is uniformly minimum variance unbiased estimator (UMVUE) by Lehmann-Scheffe's Theorem (Lehman & Casella, 1998).

Theorem 12. Let x_1, x_2, \dots, x_n be random variables distributed as fractal with parameter λ . The uniformly minimum variance unbiased estimator (UMVUE) of λ is

$$\hat{\lambda} = 1 + \frac{n-1}{n} \cdot \frac{1}{\bar{y}}$$

Proof. We prove the unbiasedness of the estimator by Theorem 9. By Theorem 10, $\hat{\lambda}$ is a function of complete and sufficient statistic. Now, we find the variance of $\hat{\lambda}$.

$$Var(\hat{\lambda}) = E(\hat{\lambda}^2) - (E(\hat{\lambda}))^2$$

Since $\hat{\lambda}$ is unbiased estimator, then $E(\hat{\lambda}) = \lambda$.

Find $E(\hat{\lambda}^2)$

$$E(\hat{\lambda}^2) = E\left[\left(1 + \frac{n-1}{n} \cdot \frac{1}{\bar{y}}\right)^2\right]$$

$$E(\hat{\lambda}^2) = E\left[1 + \frac{2(n-1)}{n} \cdot \frac{1}{\bar{y}} + \frac{(n-1)^2}{n^2} \cdot \left(\frac{1}{\bar{y}}\right)^2\right]$$

$$E(\hat{\lambda}^2) = 1 + 2(n-1)E(s^{-1}) + (n-1)^2E(s^{-2})$$

Using (10) and (12),

$$E(\hat{\lambda}^2) = 1 + 2\beta + \frac{(n-1)\beta^2}{n-2} \tag{13}$$

With $\beta = \lambda - 1$,

$$Var(\hat{\lambda}) = 1 + 2(\lambda - 1) + \frac{(n-1)(\lambda-1)^2}{n-2} - \lambda^2$$

$$Var(\hat{\lambda}) = \frac{(n-1)(\lambda-1)^2}{n-2} - (\lambda - 1)^2$$

$$Var(\hat{\lambda}) = \left(\frac{n-1}{n-2} - 1\right) (\lambda - 1)^2$$

$$Var(\hat{\lambda}) = \frac{1}{n-2} (\lambda - 1)^2 \tag{14}$$

Now, we obtained the Cramer-Rao Lower Bound for λ . The Fisher information is $I(\lambda)$:

$$\begin{aligned} L &= \prod_{i=1}^n \frac{\lambda-1}{\theta} \left(\frac{x_i}{\theta}\right)^{-\lambda} \\ \ln L &= n \ln(\lambda - 1) - n \ln \theta - \lambda \sum_{i=1}^n \ln \frac{x_i}{\theta} \\ \frac{\partial \ln L}{\partial \lambda} &= \frac{n}{\lambda-1} - \sum_{i=1}^n \ln \frac{x_i}{\theta} \\ \frac{\partial^2 \ln L}{\partial \lambda^2} &= -\frac{n}{(\lambda-1)^2} \\ I(\lambda) &= -E\left(\frac{\partial^2 \ln L}{\partial \lambda^2}\right) \\ I(\lambda) &= \frac{n}{(\lambda-1)^2} \\ Var(\hat{\lambda})_{CR} &= \frac{1}{nI(\lambda)} = \frac{(\lambda-1)^2}{n} \end{aligned} \tag{15}$$

By (14) and (15), $Var(\hat{\lambda}) > Var(\hat{\lambda})_{CR}$. However, as $n \rightarrow \infty$, $Var(\hat{\lambda}) > Var(\hat{\lambda})_{CR}$. By Lehmann-Scheffe's Theorem (Lehman & Casella, 1998), λ which is an unbiased estimator and function of sufficient and complete statistic has the minimum variance among unbiased estimators of λ . ■

6.0 Conclusion

The characteristics of fractal data are described by fractal dimension. It is extremely important to have the best possible estimate of the fractal dimension. We have shown that estimator of λ is uniform minimum unbiased estimator (UMVUE). Having such an estimate leads to a better and more accurate statistical analysis and the basis in formulating the framework of statistical inference about the fractional dimension.

References

Bauke, H. (2007). Parameter estimation for power-law distributions by maximum likelihood methods. *European Physical Journal B*, 58,167–173. <https://doi.org/10.1140/epjb/e2007-00219-y>

- Clauset, A., Shalizi, C. R., & Newman, M. E. J. (2009). Power-law distributions in empirical data. *SIAM Review*, 51(4), 661–703. <https://doi.org/10.1137/070710111>
- Feller, W. (1991). *An introduction to probability theory and its applications* (2nd ed.). New York.
- Lehman, P. (1998) *Theory of point estimation* (2nd ed.). Springer-Verlag.
- Malik, H. J. (1970). Estimation of the parameters of the PARETO distribution. *Metrika*. 15,126–132. <https://doi.org/10.1007/BF02613565>
- Mandelbrot, B. B. (1983). *The fractal geometry of nature*. W. H. Freeman.
- Mandelbrot, B.B., & Van Ness, J.W. (1968). Fractional brownian motions, fractional noises and applications. *SIAM Review*, 10(4), 422–437 <https://www.jstor.org/stable/2027184?seq=1>
- Newman, M. (2005). Power laws, Pareto distributions and Zipf's law. *Contemporary Physics*, 46(5), 323-351. <https://doi.org/10.1080/00107510500052444>
- Padua, R.N., Baldado, M., Adanza, J. G., & Panduyos, J. B. (2013), Statistical analysis of fractal observations: Applications in education and in poverty estimation. *SDSSU Multidisciplinary Research Journal*, 1(1). <https://www.smrj.sdssu.edu.ph/index.php/SMRJ/article/view/84>
- Padua, R. N., & Borres, M.S. (2013). From fractal geometry to fractal statistics. *Recoletos Multidisciplinary Journal of Research*, 1(1). <https://doi.org/10.32871/rmrj1301.01.09>
- Padua, R. N., Ontoy, D. S., Palompon, D. R., & Mirasol, J. M. (2015). Statistical fractal inference. *University of the Visayas - Journal of Research*, 9(1), 15-22. <https://doi.org/10.5281/zenodo.Z144190>
- Ramachandran, K.M., & Tsokos, C.P. (2009). *Mathematical statistics with applications*. Elsevier. <https://books.google.com.ph/books?id=YFyhXk-ONWwC&printsec=frontcover#v=onepage&q&f=false>
- Regalado, D. Y., Padua, R.N., Azura, R.B., & Perez, K. B. (2018). Minimum variance unbiased estimation of the scale parameter of exponential distributions and related logarithmic integrals. *Journal of Higher Education Research Disciplines*, 3(1). <https://nmsc.edu.ph/ojs/index.php/jherd/article/view/97>
- Ross, S.M. (2009). *Introduction to probability and statistics for engineers and scientists* (4th ed.). Academic Press. <https://doi.org/10.1016/B978-0-12-370483-2.X0001-X>
- Temme, N. M. (2010), Exponential, logarithmic, sine, and cosine integrals. In F.W. Olver, D. W. Lozier, R. F. Boisvert, & C. W. Clark (Eds.), *The NIST handbook of mathematical functions*. Cambridge University Press.