FRACTAL STATISTICAL ANALYSIS

[1]*Adriano V. Patac, Jr.** and [2]*Roberto N. Padua*

Abstract

This paper provides the statistical foundations for fractal analysis of real-life data. It begins by developing a fractal probability distribution based on a power-law formulation and proceeds by estimating the parameters through (a) maximum likelihood estimation method, (b) exponential parameter estimation, and (c) regression approach. Data analysis based on a fractal distribution hypothesis is heavily guided by the fact that for a random variable X with fractal distribution $f(x;\theta,\lambda)$, the random variable $y=\log(x/\theta)$ has an exponential distribution with rate parameter $\beta = \lambda - 1$. A new Q-Q plot is introduced for assessing the fractality of observations. Likewise, tests of hypotheses about the fractal dimension λ are introduced based on the pivotal statistic $y=\log(x/\theta)$. Test statistics are constructed whose null distributions are shown to be chi-square, for testing H_0: $\lambda=\lambda_0$ or beta distributions, for testing H_0: $\lambda_1=\lambda_2$.

Keywords: fractal statistical dimension, fractal random variable, power-law distribution

*Corresponding Author: Adriano V. Patac Jr., adriano.patac@gmail.com

1.0 Introduction

Since the publication of Mandelbrot's (1983) book entitled "The Fractal Geometry of Nature", interest in this science has grown exponentially. Applications of fractal geometry are noted in various fields of inquiry: medicine, ecology, agriculture, communications, engineering and the sciences (Palmer, 1988; Cohen, 1997; Russel, 1995) with surprising diversity. The full potential of fractal geometry can be more effectively achieved if analytic machinery can be developed to analyze numerical observations. This paper attempts to develop fractal statistics for this purpose using a power law formulation for fractal random variables.

The power law distribution is found to be pervasive in a wide variety of physical, biological, and man-made phenomena. These include the magnitudes of seismic tremors, the variety of moon crater sizes and of solar flares (Neukum & Ivanov, 1994), the food-search pattern of various animals (Humphries et al., 2010), the sizes of activity patterns of neuronal populations (Klaus et al., 2011), the occurrence selected words in many languages, frequencies of family names, the species richness in clades of organisms (Albert et al., 2011), the sizes of power outages, wars, criminal charges per convict, and many other quantities (Newman, 2005). Few empirical distributions fit a power law for all their values, but rather follow a power law in the tail.

A central concept in fractal geometry is the notion of a fractal dimension (λ). In the usual Euclidean geometry, the dimension of a geometric object is a non-negative integer less than or equal to 3. Thus, a point has zero (0) dimension; a line has one (1) dimension; a plane has two (2) dimensions and a cube has three dimensions. The generalization of the usual Euclidean geometry to accommodate dimensions greater than 3, namely, a Hilbert space, maintains the restriction of a positive integer dimension for the geometric entities in the space. Mandelbrot (1983) posited that geometric objects possessing non-integer dimensions can be constructed viz. on a line, a geometric object can be constructed that has dimension between 0 and 1, like 0.63, provided that the term "dimension" is properly defined. The fractal dimension of a geometric object is characterized as its space-filling property.

Fractals are viewed as rugged, irregularly-shaped geometric objects arising out of endless repetitions of the same pattern at all scales. Self-similarly and scale-invariance characterize all fractals and because of these, the box-counting dimension is defined as:

$$1.)\quad \lambda = \frac{\log(m)}{\log(r)}$$

where m = number of copies itself and r = scaling factor. The Cantor set or fractal dust has λ = 0.63; the Von Koch curve has λ= 1.26; the Sierpinski carpet has λ= 1.89. the higher the fractal dimension of an object, the more rugged it is compared to its Euclidean version. Thus, fractal dusts are coarser than points; the Von Koch curve is rougher than a line and so on.

To translate these purely geometric notions in statistical terms requires the development of a probabilistic model that captures the ideas of self-similarity and scale-invariance leading to an observed randomness or data roughness. Section 2 develops the concept of a statistical fractal distribution. Section 3 concentrates on parameter estimation for the statistical model. Finally, Section 4 provides statistical methods for testing hypotheses about the fractal dimension of random variables.

2.0 Model Development

The idea of self-similarity and scale invariance pervades the development of a probabilistic model for fractal observations. One attribute of power laws is their scale invariance. Given a relation $f(x)=ax^k$, scaling the argument x by a constant factor c causes only a proportionate scaling of the function itself. That is,

$$f(cx) = a(cx)^k = ck\, f(x) \propto f(x)$$

[2]*Surigao State College of Science and Technology*
[1]*Liceo de Cagayan University*

That is, scaling by a constant c simply multiplies the original power-law relation by the constant c^k. Thus, it follows that all power laws with a particular scaling exponent are equivalent up to constant factors, since each is simply a scaled version of the others.

Definition: Let X>0 be a positive random variable with distribution function $F(x)$. Suppose that $F'(x) = f(x)$ exists and is continuous for all positive x. Let $f(x) = ax^{-\lambda}, \lambda > 0, a > 0, x > \theta > 0$. Then, X is called a fractal random variable.

We note that the density function $f(x)$ puts large mass for small values of x. Moreover, it is self-similar and scale-invariant because $f(cx)=a(cx)^{-\lambda} = c^{-\lambda}(ax^{-\lambda})=c^{-\lambda} f(x)$. The constant $\boldsymbol{a}$ is completely determined by the parameters θ and λ; since

$$\int_{\theta}^{\infty} f(x) = \int_{\theta}^{\infty} ax^{-\lambda}dx = a\frac{x^{1-\lambda}}{1-\lambda}\Big|_{\theta}^{\infty} = 1$$

and:

$$a = \frac{\lambda - 1}{\theta^{1-\lambda}}$$

Hence:

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

The definition of $f(x)$ in (2) implies that for a fractal distribution:

3.) $f(x) \propto \left(\frac{x}{\theta}\right)^{-\lambda}$

from which:

4.) $-\lambda \propto \frac{\log f(x)}{\log\left(\frac{x}{\theta}\right)}.$

Recalling that the box-counting dimension of a fractal object is given by (1), we make the association $m = f(x)$ and $r = \left(\frac{\theta}{x}\right)$ from (4), and

Definition: The exponent λ in the fractal distribution $f(x;\lambda,\theta)$ is called the fractal dimension of the random variable X. The probability density function $f(x)$ determine the "number of copies" observed for x over the scale $\frac{\theta}{x}$.

Let $X>\theta$ have the fractal distribution (2), Then:

Theorem 1. Let $f(x) = \left(\frac{\lambda-1}{\theta}\right)\left(\frac{x}{\theta}\right)^{-\lambda}, \lambda > 1, x > \theta > 0.$ Then $y = log\left(\frac{x}{\theta}\right)$ has an exponential distribution with parameter $\beta=\lambda-1$.

Proof: Let $y = log\left(\frac{x}{\theta}\right)$, then $\frac{dx}{dy} = \theta e^{y}$. Using transformation of variables, we obtain:

$$g(y) = (\lambda - 1)e^{-(\lambda-1)y}, y > 0. \blacksquare$$

For an exponential random variable y with parameter β, we have that:

5.) $E(y) = \frac{1}{\beta}.$

Equation (5) is useful for estimating the fractal dimension λ. Let $y_i = log\left(\frac{x_i}{\theta}\right), i = 1, 2, \ldots, n,$ then $\bar{y} = \frac{\sum_{i=1}^{n} log\left(\frac{x_i}{\theta}\right)}{n}$ is an estimate of $E(y)$. Since $\beta=\lambda-1$, it follows from (5) that an estimate for λ can be obtained from:

6.) $\bar{y} = \frac{1}{\hat{\lambda}-1}$ or $\hat{\lambda} = 1 + \frac{1}{\bar{y}} = 1 + \frac{n}{\sum_{i=1}^{n} log\left(\frac{x_i}{\theta}\right)}.$

3.0 Estimation of Parameters

The cumulative distribution function $F(x)$ of the random variable x is given by:

7.) $\ldots F(x) = 1 - \left(\frac{x}{\theta}\right)^{1-\lambda}, \lambda > 1, \theta > 0$

which contains the parameters λ and θ. It is easy to show that the maximum-likelihood estimators of λ and θ

$\hat{\lambda} = 1 + \frac{n}{\sum_{i=1}^{n} log\left(\frac{x_i}{\theta}\right)} = 1 + \frac{1}{\bar{y}}$ where $\bar{y} = \frac{1}{n}\sum_{i=1}^{n} log\left(\frac{x_i}{\theta}\right)$

$\hat{\theta} = \min_i\{x_i\}.$ respectively are:

8.)

The maximum-likelihood estimator (MLE) of λ coincides with (6) which was derived using the fact that $y = log\left(\frac{x}{\theta}\right)$ $\log f(x) = \log a + -\lambda \log\left(\frac{x}{\theta}\right)$ where $a = \frac{\lambda-1}{\theta}$. eter (λ-1).

A slightly different estimator of λ can be derived using (4). From (4) we deduce that:

9.) $y_i^* = \beta_0 + \beta_1 x_i + \varepsilon_i, i = 1, 2, \ldots, n$ where $\beta_0 = \log a,\ x_i^* = log\left(\frac{x}{\theta}\right)$

which is a linear model of the form:

10.) $\forall i.$ and

$y_i^* = log f(x_i), \beta_1 = -\lambda$

The err $\hat{\beta}_1 = \frac{n\sum_{i=1}^{n}\left[log\left(\frac{x_i}{\theta}\right)\right][\log f(x_i)] - \sum_{i=1}^{n} log\left(\frac{x_i}{\theta}\right)\sum_{i=1}^{n}\log f(x_i)}{n\sum_{i=1}^{n}\left(log\left(\frac{x_i}{\theta}\right)\right)^2 - \left[\sum log\left(\frac{x_i}{\theta}\right)\right]^2} = -\hat{\lambda}.$ l identically di estimator for λ is $-\hat{\beta}_1$:

11.) $\hat{\beta}_1$

Theorem 2. For the linear model (10), the least-squares estimator of β_1 is an unbiased estimator of –λ.

Proof: From linear models, $E(\hat{\beta}_1)=\beta_1$ (Graybill, 1976) and

$$Var(\hat{\beta}_1) = \frac{1}{\sum\left[log\left(\frac{x_i}{\theta}\right) - \frac{1}{n}\sum_{i=1}^{n} log\left(\frac{x_i}{\theta}\right)\right]^2} = \frac{1}{\sum_{i=1}^{n}(y_i - \bar{y})^2}$$

The s. From linear mo $y_i = log\left(\frac{x_i}{\theta}\right)$ ve: $\bar{y} = \frac{1}{n}\sum_{i=1}^{n} log\left(\frac{x_i}{\theta}\right)$

12.) , where

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

and

Since $y_i \underset{\sim}{d}\, exponential\,(\lambda-1)$, it follows that $Var(y_i) = \frac{1}{(\lambda-1)^2}$ and so:

13.) for large n.

Theorem 3. Let $\hat{\beta}_1$ be given by Equation (11). Then:

$$\sqrt{n}\,(\lambda-\hat{\lambda}) \xrightarrow{L} N(0,(\lambda-1)^2\,).).$$

Proof: Apply the central limit theorem and equation (13). Observe that Theorem 3 is equivalent to:

The estimator (11) of λ is the regression estimator which differs from the maximum likelihood estimator (8). In order to study the behavior of the MLE of λ, we need to observe that the strong law of large numbers(SLLN) hold for:

$$S_n = \sum_{i=1}^{n} log\left(\frac{x_i}{\theta}\right)$$

That is: $z_n = \frac{S_n}{n} \to E\left(log\left(\frac{x_i}{\theta}\right)\right) = \frac{1}{\beta}$ with probability 1 (wp1).

Slutsky's lemma states that "if $z_n \to \alpha$ wp1 and $g(x)$ is a continuous function continuous at x=α, then $g(z_n) \to g(x)$wp1." To make use of this lemma, let $g(x) = 1 + \frac{1}{x}, x \neq 0.$ Then, $g(x)$ is a continuous function of x. Likewise:

$$g(z_n) = 1 + \frac{1}{z_n} = 1 + \frac{1}{\bar{y}} = 1 + \frac{n}{\sum_{i=1}^{n} log(x_i)} = \hat{\lambda}$$

$$\hat{\lambda} = g(z_n) \to 1 + \frac{1}{\frac{1}{\beta}} = 1 + \beta = \lambda\, wp1$$

However, $(z_n) \to \frac{1}{\beta} = \frac{1}{\lambda-1}$ with probability 1, hence:

We have essentially proved that:

Theorem 4. The maximum likelihood estimator $\hat{\lambda}$ given by equation (8) converges to λ with probability 1.

Since convergence with probability 1 implies convergence in probability i.e SLLN implies WLLN, Theorem 4 implies Theorem 3, Unlike the regression estimator of λ, however, the MLE of λ can be *biased* since:

$$E\left(1 + \frac{1}{E(\bar{y})}\right) \neq 1 + \frac{1}{E(\bar{y})} = \lambda$$

In fact, by Jensen's inequality upon noting that $\varphi(x)$=$1 + \frac{1}{x}$ is a convex function of x, we obtain:

$$1 + \frac{1}{E(\bar{y})} = \lambda \;\leq\; E\left(1 + \frac{1}{\bar{y}}\right)$$

It is not difficult to show that the bias of the maximum likelihood estimator is of the order $O(n^{-1})$ which is rather small for n>100.

Similarly, Jensen's inequality also implies:

14.) $Var\left(\hat{\lambda}_{MLE}\right) = Var\left(1 + \frac{1}{\bar{y}}\right) = Var\left(\frac{1}{\bar{y}}\right) \geq \frac{1}{Var(\bar{y})} = \frac{(\lambda-1)^2}{n}.$

Equation (14) means that the uncertainty in maximum likelihood estimation is given by

$$\sigma = \frac{\lambda-1}{\sqrt{n}}$$

In summary, we found that the maximum likelihood estimator of λ is both asymptotically unbiased and efficient for λ. In contrast, the regression estimator of λ is unbiased and asymptotically efficient as well.

4.0 Data Analysis

The fractal distribution represented by Equation (2) have a properly defined mean only if the fractal dimension λ exceeds 2 and will have a finite variance when the fractal dimension exceeds 3. In practical situations, only the mean exists but the variance may not. Most identified power laws in nature have exponents such that the mean is well-defined but the variance is not, implying they are capable of having very large means due to a few large observations. A typical example is income. The distribution of incomes is typically right-skewed in that more people have lower incomes than higher incomes. The presence of just one millionaire in the group can easily distort the average income of a group. Income is distributed according to a power-law known as the Pareto distribution.

On the one hand, the use of traditional statistics to handle cases where the mean and the variance of a random variable X are assumed to exist will lead to misleading conclusions. These include the usual analysis of variance and regression approaches. Meanwhile, fractal statistics allows for more efficient interventions not otherwise available in traditional normal-based statistics. For example, if we know that the exhausts from cars are distributed according to some power law i.e. only very few cars contributed largely to the air pollution, then it would be sufficient to eliminate those very few cars from the road to reduce total exhaust substantially.

4.1 Assessing Fractal Characteristics From Data

The pivotal statistics that will be used throughout the analysis of fractal observations is the statistic:

$$y = log\left(\frac{x}{\theta}\right),\;\; \theta = \min\{x_i\}$$

Let $x_1, x_2, \ldots, x_n$ be a random sample from a distribution $F(x;\, \lambda,\, \theta)$, x >0, assumed absolutely continuous with respect to a Lebesque measure. Then, the statistic $y_i = log\left(\frac{x_i}{\theta}\right)$ has an exponential distribution with parameter $\beta = \lambda$-1 if and only if X comes from the distribution:

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

We have proved the "if" part. We now demonstrate the converse. Since $y_i=log\left(\frac{x_i}{\theta}\right) \underset{\sim}{d}\ exponential\ (\beta=\lambda-1)$, then:

$$g(y)=(\lambda\text{-}1)\ e^{-(\lambda-1)y},\quad y>0$$

The Jacobian of this transformation is: $J=\left|\frac{dy}{dx}\right|=\frac{1}{x}$. Thus:

$$g(x)=(\lambda-1)e^{-(\lambda-1)\log\left(\frac{x}{\theta}\right)}.\frac{1}{x}=\frac{(\lambda-1)}{\theta}\left(\frac{x}{\theta}\right)^{-\lambda},\lambda>1,x>\theta.$$

Consider the transformed data $y_1,y_2,...,y_n$. if the observations $x_1, x_2,..., x_n$ are iid $f(x;\lambda,\theta)$, then $y_1,y_2,..,y_n$. are $iid\ exponential$ $(\beta=\lambda\text{-}1)$. Let $\bar{y}=\frac{1}{n}\sum_{i=1}^{n}y_i$, then $E(\bar{y})=\frac{1}{\beta}$ so that $\hat{\beta}=\frac{1}{\bar{y}}$ is an estimate of the rate parameter β of the exponential distribution. A plot of the theoretical quantities of the exponential distribution with parameter β with the observed quantiles $y_{(1)}\leq y_{(2)}\leq...\leq y_{(n)}$ will yield a straight line with positive slope (m=1). A formal test of hypothesis can be carried out for this purpose. The theoretical α^{th} quantile from an exponential distribution with parameter β is:

15.) $x_\alpha=\frac{-\ln(1-\alpha)}{\beta}\quad \alpha=\frac{1}{n+1},\frac{2}{n+1},\frac{3}{n+1},\ldots,\frac{n}{n+1}$

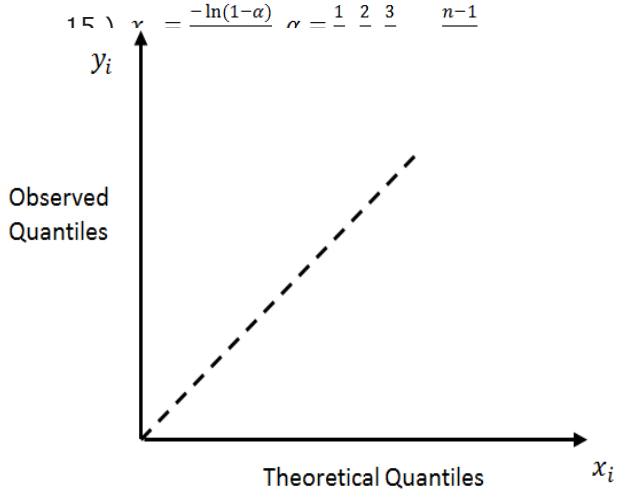

Figure 1: Q-Q plot of the theoretical quantiles of $exp(\beta)$ and the observed quantiles of $y_i=log\left(\frac{x_i}{\theta}\right)$.

A necessary condition for a random variable X to possess a fractal distribution $f(x;\theta,\lambda)$ is for it to be right-skewed. That is, if $\mu=E(x)$ exists, then $\mu>\tilde{\mu}$ where $\tilde{\mu}=median\ (x)$. The expression for μ and $\tilde{\mu}$ are:

16.) $\mu=E(x)=\left(\frac{\lambda-1}{\lambda-2}\right)\theta$

$$\left(\frac{\lambda-1}{\lambda-2}\right)^{\frac{1}{1-\lambda}}>\frac{1}{2},$$

and the condition μ and $\tilde{\mu}$ can be expressed as:

17.) for $\lambda > 2$.

4.2. Test of Hypothesis About λ

Since the distribution of $y_i=log\left(\frac{x_i}{\theta}\right)$ is known to be exponential with rate parameter $\beta=\lambda$-1, then a test of hypothesis about λ is equivalent to a test of hypothesis about β. Consider a simple test of hypothesis about λ:

H_0: $\lambda=\lambda_0$ versus H_1: $\lambda=\lambda_1$, $\quad\lambda_1<\lambda_0$

Or equivalently:

H_0: $\beta=\beta_0$ versus H_1: $\beta=\beta_1$, $\quad 0<\beta_1<\beta_0$

Let $x_1, x_2,...,x_n$ be a random sample from $f(x;\lambda,\theta)$. Then, $y_1, y_2,...,y_n$ is a random sample from an exponential distribution with $\beta=\lambda$-1 and $y_i=log\left(\frac{x_i}{\theta}\right)$. Assume θ is known. The likelihood function $l(y_1,...,y_n|\beta_0)$ is:

18.) $l(y_1,...,y_n|\beta_0)=\beta_0^n\ exp(-\beta_0\sum_{i=1}^{n}y_i).$

Similarly, the likelihood functions under H_1 is:

19.) $l(y_1,...,y_n|\beta_1)=\beta_1^n\ exp(-\beta_1\sum_{i=1}^{n}y_i).$

The likelihood ration (LR) is thus:

20.) $LR=\left(\frac{\beta_0}{\beta_1}\right)^n\ exp\left((\beta_1-\beta_0)\sum_{i=1}^{n}y_i\right)$

If the null hypothesis were false, then $LR\leq k$ for some constant k with probability α. That is:

21.) $$\alpha=P(LR\leq k)$$

$$=P\left[\left(\frac{\beta_0}{\beta_1}\right)^n\exp(\beta_1-\beta_0)\sum_{i=1}^{n}y_i\leq k\right]$$

$$=P[(\beta_1-\beta_0)\sum y_i\leq\ln k+n\ln\beta_1+n\ln\beta_0]$$

$$=P\left[\sum y_i\leq\frac{\ln k+n\ln\beta_1+n\ln\beta_0}{(\beta_1-\beta_0)}\right]$$

Let $V=2\beta_0\sum_{i=1}^{n}y_i=2\beta_0\sum_{i=1}^{n}log\left(\frac{x_i}{\theta}\right)$. Then it is easy to show that $V\ \underset{\sim}{d}\ \chi^2_{2n}$ under H_0. Thus, reject H_0 iff $V\leq\chi^2_\alpha(2n)$.

4.3 Testing the Equality of Two (2) Fractal Dimensions.

Let $x_1,x_2,...,x_{n_1}$ be $iid\ f(x;\lambda_1,\theta)$ and $y_1,y_2,...,y_{n_2}$ be $iid\ f(y;\lambda_2,\theta)$. The statistics $\left\{u_i=log\left(\frac{x_i}{\theta}\right)\right\}$ and $\left\{v_i=log\left(\frac{y_i}{\theta}\right)\right\}$ $i=1,2,3,...,n$ are then iid random samples from $exp(\beta_1=\lambda_1\text{-}1)$ and $exp(\beta_2=\lambda_2\text{-}2)$ respectively. We want to test H_0: $\lambda_1=\lambda_2$ versus H_1: $\lambda_1\neq\lambda_2$ at α level or equivalently: H_0: $\beta_1=\beta_2$ versus H_1: $\beta_1\neq\beta_2$ We state a result due to Nydick (2012):

Result 1: Let μ and v be independent exponential variates with the same rate parameters λ. Then, $x=\frac{\mu}{v}$ is distributed as:

$$f_x(x)=\frac{1}{(1+x)^2}\quad,x>0.$$

Let $\mu = \sum_{i=1}^{n_1} log\left(\frac{x_i}{\theta}\right)$ and $v = \sum_{i=1}^{n_2} log\left(\frac{y_i}{\theta}\right)$. Using the moment-generating function technique, it is easy to show that:

22.) $\mu \underset{\sim}{d}\ Gamma(n_1, \beta_1)$
$v \underset{\sim}{d}\ Gamma(n_2, \beta_2).$

We extended the result of Nydick (2012) to:

Result 2: (Generalization) Let $\mu \underset{\sim}{d}\ Gamma(n_1, \beta_1)$ and $v \underset{\sim}{d}\ Gamma(n_2, \beta_2)$. If $\beta_1 = \beta_2$, then the distribution of $z = \frac{u}{v}$ is:

$$f_z(z) = \frac{\Gamma(n_1+n_2)}{\Gamma(n_1)\Gamma(n_2)} \frac{z^{n_1-1}}{(1+z)^{n_1+n_2}} = \frac{1}{\beta(n_1+n_2)} \frac{z^{n_1-1}}{(1+z)^{n_1+n_2}}$$

where $\beta(n_1, n_2)$ is the beta function of the second kind:

$$\beta(n_1, n_2) = \int_0^1 x^{n_1}(1-x)^{n_1-1}dx = \int_0^\infty \frac{z^{n_1-1}}{(1+z)^{n_1+n_2}}dz.$$

Proof:

Let $u \underset{\sim}{d}\ Gamma(n_1, \beta)$ and $v \underset{\sim}{d}\ Gamma(n_2, \beta)$. Since μ and v are independent, it follows that their joint distribution is given by:

$$f(\mu, v) = \frac{\beta^{n_1+n_2}}{\Gamma(n_1)\Gamma(n_2)} \mu^{n_1-1} v^{n_2-1} \exp(-\beta(\mu+v))$$

Let $z = \frac{\mu}{v}$ and $w = v$. It follows that $\mu = wv$ and $v = w$, from which the Jacobian of the transformation is found to be $J = w$. The joint distribution of z and w becomes:

$$f(wz, w) = \frac{\beta^{n_1+n_2}}{\Gamma(n_1)\Gamma(n_2)} (wz)^{n_1-1} w^{n_2-1} \exp(-\beta(1+z)w)J$$

which simplifies to:

$$f(wz, w) = \frac{\beta^{n_1+n_2}}{\Gamma(n_1)\Gamma(n_2)} z^{n_1-1} w^{n_1+n_2-1} \exp(-\beta(1+z)w)$$

To find the marginal distribution of z, we integrate out w, to obtain:

$$f(z) = \frac{\Gamma(n_1+n_2)}{\Gamma(n_1)\Gamma(n_2)} \frac{z^{n_1-1}}{(1+z)^{n_1+n_2}} = \frac{1}{\beta(n_1+n_2)} \frac{z^{n_1-1}}{(1+z)^{n_1+n_2}}$$

where $\beta(n_1, n_2)$ is the beta function of the second kind:

$$\beta(n_1, n_2) = \int_0^1 x^{n_1}(1-x)^{n_1-1}dx = \int_0^\infty \frac{z^{n_1-1}}{(1+z)^{n_1+n_2}}dz. \blacksquare$$

Thus, for testing H_0: $\lambda_1 = \lambda_2$ versus H_1: $\lambda_1 \neq \lambda_2$ at α level, the decision rule is to reject H_0 iff $= \frac{\mu}{v} > \beta_\alpha(n_1, n_2)$. In practice, it is more convenient to refer to the F distribution rather than the beta distribution. This can be achieved through the following well-known relationship between the beta distribution and the F-distribution.

Result 3: If $z \underset{\sim}{d}\ Beta(\alpha_1, \alpha_2)$, then $f = \frac{\alpha_2 z}{\alpha_1(1-z)}$ has an F-distribution, $F(2\alpha_1, 2\alpha_2)$.

The decision criterion now becomes: "Reject H_0 iff $f = \frac{n_2 z}{n_1(1-z)} > F_\alpha(2n_1, 2n_2)$".

References

Albert, J.S. & Reis, R.E., eds. (2011). *Historical biogeography of neotropical freshwater fishes*. Berkeley: University of California Press.388pp.

Cohen, N. (1997 May). *Fractalstorical antenna application in wireless telecommunication. In Electronics Industries Forum of New England,* 1997. Professional Program Proceedings (pp. 43-49). IEEE.

Graybill, F. A. (1976). Theory and application of the linear model. Duxbury, North Scituate, Massachusetts.

Humphries, N.E., Queiroz, N., Dyer, J.R., Pade, N.G., Musyl, M.K., Schaefer, K.M., Fuller, D.W., Brunnschweiler, J.M., Doyle, T.K., Houghton, J.D., Hays, G.C., Jones, C.S., Noble, L.R., Wearmouth, V.J., Southall, E.J., Sims, D.W. (2010). Environmental context explains Lévy and Brownian movement patterns of marine predators. *Nature, 465 (7301), 1066–1069.* doi:10.1038/nature09116

Klaus, A., Yu, S. & Plenz, D. (2011). Statistical analyses support power law distributions found in neuronal avalanches. *PLoS ONE, 6(5), e19779*. doi:10.1371/journal.pone.0019779

Mandelbrot, B.B. (1983). *The fractal geometry of nature*. W.H. Freeman. San Francisco. California.

Neukum, G. & Ivanov, B.A. (1994). *Crater size distributions and impact probabilities on Earth from lunar, terrestrial-planet, and asteroid cratering data.* In T. Gehrels (ed.), Hazards Due to Comets and Asteroids, University of Arizona Press, Tucson, Arizona, 359–416.

Newman, M.E. J. (2005). Power laws, pareto distributions and Zipf's law". *Contemporary Physics, 46(5): 323–351.*

Nydick, S.W. (2012). *A different(ial) way matrix derivatives again.* University of Minnesota. Retrieve from http://www.tc.umn.edu/~nydic001/docs/unpubs/Magnus_Matrix_Differentials_Presentation.pdf on Dec. 5, 2014.

Russell, B.D., Mamishev, A.V., & Benner, C.L. (1994, mag.) *Analysis of lign impedance faults using fractal techniques. In Power Industry Computer Application Conference,* 1995. Conference Proceeding, 1995 IEEE (pp.401-406). IEEE

Palmer, M.W. (1988). Fractal geometry: a tool for describing spatial patterns of plant communities. *Vegetatio, 75 (1), 91–102*. doi:10.1007/BF00044631