APPLICATIONS OF A BIVARIATE EXPONENTIAL DISTRIBUTION WITH LINEAR STRUCTURES IN FRACTAL STATISTICS

[1]*Roberto N. Padua*, [2]*Vincent T. Lapinig and* [3]*Johnny Mark M. Bolante*
[1]Liceo de Cagayan University, [2]Governor Alfonso D. Tan College,
[3]Northwestern Mindanao State College of Science and Technology

ABSTRACT

This paper explores the properties of a bivariate exponential distribution with linear structures where the Pearson product moment correlation exists as proposed by Iyer et al. (2002). These properties were then used as the basis for constructing a Pearson-type correlation of the underlying fractal random variables obtained from the marginal distributions of the bivariate exponential distribution. A link function h (.) which defines the relationship between the correlation values obtained from the bivariate exponential distribution and the correlation values obtained from the bivariate fractal distribution whose fractal dimensions exceed 3 was obtained. The results were then applied to analyze the Philippine data on typhoons, casualties, and damage from 1960 to 2013. The results obtained by applying the link function h (.) revealed that while the estimated damage (D) and number of casualties (C) correlated significantly, the corresponding economic impact (X) and number of people affected (Y) were not. In order to link the latter two variables, information on the characteristics of the casualties' viz. breadwinner or not, are needed.

Keywords: copula, bivariate exponential, bivariate fractal, fractal correlation

Introduction

The fundamental theorem in fractal statistics states that a random variable X has a power law distribution if and only if $log\left(\frac{x}{\theta}\right)$, θ where is the minimum value of X, has an exponential distribution with rate parameter $\beta = \lambda - 1$ and λ is the fractal dimension of X (Padua, 2014). The fundamental theorem clearly connects the power-law distribution with an exponential distribution, so that if the exponential distribution is given, it is possible to deduce the corresponding power-law distribution. In the bivariate case, if a bivariate exponential distribution with linear structures having exponential marginals is given, then the Pearson product-moment correlation of the exponential random variables X and Y can be used to deduce the correlation between the underlying fractal random variables z_1 and z_2 if this exists.

More specifically, suppose that X and Y have a bivariate exponential distribution $f(x, y)$ with Pearson correlation of r_{xy}, then if

$$1. z_1 = \exp(x), \quad x > 0$$

$$z_1 = \exp(y), \quad y > 0$$

the joint distribution $h(z_1, z_2)$ has a power-law form with correlation r_{z_1, z_2} provided that $\lambda_1 > 3$ and $\lambda_2 > 3$. There are formulations of the bivariate exponential distribution which allow for the existence of the Pearson product-moment correlation.

Iyer et al. (2002) proposed a bivariate exponential distribution with linear structures that has a Pearson correlation r_{xy} expressible as a function of the dependence parameter ρ. The other proposed bivariate exponential distributions do not possess this property. For instance, the Marshall-Olkin bivariate distribution:

$$2. f(x, y) = \exp(-\lambda_1 x - \lambda_2 y - \lambda_{12} \max(x, y)), \ x > 0, \ y > 0$$

has no representation for the Pearson correlation coefficient (Marshall and Olkin (1967)). Kumar (2010) expounded on a bivariate exponential distribution using the Ali-Mikhail-Haq copula:

$$3. c(u, v) = \frac{uv}{1 - \theta(1-u)(1-v)} \qquad , \theta \in [-1, 1]$$

where u and ν are uniform on [0, 1] with dependence parameter θ (Ali et al (1978)). Replacing u by $F_1(x) = 1 - e^{-\beta_1 x}, \beta_1 > 0$ and ν by $F_2(y) = 1 - e^{-\beta_2 y}, \beta_2 > 0$, one obtains the joint distribution of x and y as:

$$4. F(x, y) = \frac{(1 - e^{-\beta_1 x})(1 - e^{-\beta_2 y})}{1 - \theta\, e^{-\beta_1 x - \beta_2 y}}, \beta_1, \beta_2 > 0$$

Again, the Pearson correlation of x and y cannot be expressed in terms of this copula. However, Kendall's tau correlation is given by:

$$5.\ \tau = \frac{3\theta-2}{3\theta} - \frac{2(1-\theta)^2 \ln(1-\theta)}{3\theta^2}.$$

In this paper, we explore the properties of a bivariate exponential distribution with linear structures and corresponding Pearson correlation coefficient as proposed by Iyer et al. (2002). We intend to use those properties to determine the Pearson correlation coefficient of the underlying fractal random variables when the fractal dimensions are greater than 3. Section 2 reviews the formulation of Iyer et al. (2002). Section 3 discusses a simulation experiment to produce a function $h(.)$ that relates the correlation r_{xy} and $r_{z_1z_2}$ viz. $r_{z_1z_2} = h(r_{xy})$. Section 4 gives an actual application of the concept to the analysis of the estimated damage (D) and number of casualties (C) due to the onset of typhoons and tropical cyclones in the Philippines from 1960 to 2013.

Fractal Random Variables and Bivariate Exponential Distributions

Let v_1 and v_2 be marginally distributed as:

$$6.\ f_i(v_i) = \frac{(\lambda_i-1)}{\theta_i}\left(\frac{v_i}{\theta_i}\right)^{-\lambda_i}, \quad \lambda_i > 3, \theta_i > 0, i = 1,2$$

with joint distribution $F(v_1, v_2)$. Padua (2014) calls these as fractal random variables. It is shown in Padua (2014) that:

$$7.\ x = log\left(\frac{v_1}{\theta_1}\right)$$

$$y = log\left(\frac{v_2}{\theta_2}\right)$$

are each marginally distributed as exponentials with rate parameters $\beta_1 = \lambda_1 - 1$ and $\beta_2 = \lambda_2 - 1$ respectively. Iyer et al. (2002) constructs the joint exponential bivariate distribution of x and y as follows:

Let $\rho = \frac{\beta_2}{\beta_1} < 1$ and define:

$$8.\ y = X + IZ$$

where I is a Bernoulli random variable such that $P_r(I = 1) = 1 - \rho$ and $P_r(I = 0) = 1 - \rho$, Z is an exponential random variable independent of X. the Laplace-Stieltjes transform of (8) is then:

9. $Y(s) = X(s).Z(s)$

from which:

10. $Z(s) = \frac{Y(s)}{X(s)} = \frac{\beta_2}{S+\beta_2}.\frac{S+\beta_1}{\beta_1} = (1-\rho)\frac{\beta_2}{S+\beta_2} + \rho.$

Equation (10) defines the distribution of Z as:

11. $F(z) = \rho + (1-\rho)\left(1 - e^{-\beta_2 Z}\right), \ z > 0,$

whose mean is $E(z) = \frac{1-\rho}{\beta_2}$. The joint distribution of X and Y has the Laplace transform:

12. $XY(s_1, s_{2)} = E(e^{-s_1 x - s_2 y})$

$$= E\left(e^{-s_1 x - s_2 (x+z)}\right)$$

$$= E\left(e^{-(s_1+s_2)x}\right).E(e^{-s_2 z})$$

$$= \left(\frac{\beta_1}{s_1+s_2+\beta_1}\right)\left(\rho + (1-\rho)\frac{\beta_2}{S+\beta_2}\right).$$

The exact form of the joint distribution $F(x, y)$ can be easily obtained by inverting (12). However, the variances and covariance of X and Y can be computed without knowledge of the explicit form of $F(x, y)$:

13. $E(xy) = E(x(x+z))$

$$= E(x^2) + E(x).E(z)$$

$$= \frac{2}{\beta_1^2} + \frac{1}{\beta_1}.\frac{(1-\rho)}{\beta_2}$$

$$cov = E(xy) - E(x)E(y)$$

$$= \frac{1}{\beta_1^2}$$

$$14.\ Var\ (x) = \frac{1}{\beta_1^2}$$

$$Var\ (y) = Var(X + IZ)$$

$$= Var\ (x) + Var(I).Var(z)$$

$$= \frac{1}{\beta_1^2} + \frac{\rho(1-\rho)}{\beta_2^2}$$

Thus:

$$15.\ r_{xy} = \frac{cov(x,y)}{\sigma_x \sigma_y} = \sqrt{\rho}.$$

Empirical Relationship of r_{xy} and r_{v1v2}.

From the correlation of the exponential variates X and Y, it is desired to estimate the correlation r_{v1v2} of the underlying fractal variates v_1 and v_2. It is, at once, clear that:

$$16.\ E(XY) = E(\log(\frac{v_1}{\theta_1})\log(\frac{v_2}{\theta_2}) \neq E((\frac{v_1}{\theta_1})(\frac{v_2}{\theta_2}) = E(e^X e^Y).$$

It follows that:

$$17.\ r_{XY} \neq r_{V1V2}$$

But:

$$18. r_{V1v2} = h(r_{XY})$$

For some function h (.). In order to estimate h (.), we performed a simulation experiment as follows:

Using Equation (8) of Iyer et al. (2002), we simulated 10,000 exponential variates X and Y with dependence parameters $\rho = 0.10\ to\ 0.90\ and\ \beta_1 = 20, \beta_2 = 2,3,4, \ldots, 18$. The resulting r_{XY} for each run is noted with the corresponding p-values. Since

$\frac{v_1}{\theta_1} = e^X$ and $\frac{v_2}{\theta_2} = e^Y$

are distributed as fractals with dimensions $\lambda_1 = \beta_1 + 1$ *and* $\lambda_2 = \beta_2 + 1$, it follows that we can compute $r_{v1v2} = r_{\exp(X)\exp(Y)}$. The pairs (r_{XY}, r_{v1v2}) were then used to produce the function h (.).

Results. Table 1 shows the results of the simulation over 10,000 runs.

Table 1. Exponential and Fractal Pearson Product Moment Correlation

Dependence Parameter($\boldsymbol{\rho}$)	**r(X,Y)**	**r(v1,v2)**
.10	.240	0.000
.15	.273	0.001
.20	.318	0.015
.25	.345	0.111
.30	.384	0.175
.35	.407	0.301
.40	.461	0.357
.45	.476	0.354
.50	.515	0.420
.55	.522	0.321
.60	.547	0.388
.65	.562	0.343
.70	.586	0.423
.75	.613	0.345
.80	.640	0.364
.85	.650	0.333
.90	.670	0.285

The results of the regression analysis are shown below:

(18) The regression equation is
r(v1,v2) = 0.584 + 0.412 log(r(x,y)

Predictor	**Coef**	**SE Coef**	**T**	**P**
Constant	0.58446	0.04805	12.16	0.000
log(r(x,	0.41211	0.05799	7.11	0.000

S = 0.07272   R-Sq = 77.1%   R-Sq(adj) = 75.6%

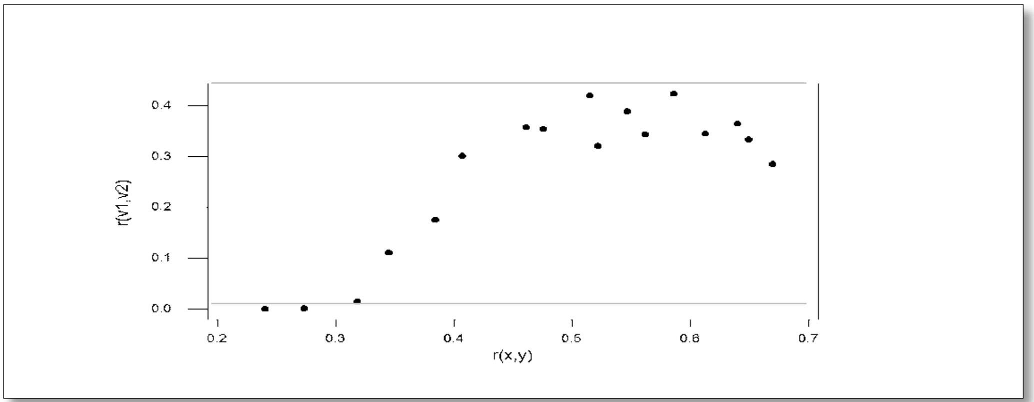

Figure 1. Scatterplot of r(x,y) versus r(v1,v2)

Application

Data gathered on the estimated damage (in million dollars) and the number of casualties reported for the various typhoons exceeding 150 kph intensity was obtained from the website of the Department of Science and Technology (DOST) Weather Bureau for the year 1960 to 2013. This information was analyzed through fractal statistics.

Figures 2 and 3 show the histogram of the number of casualties and estimated damage respectively.

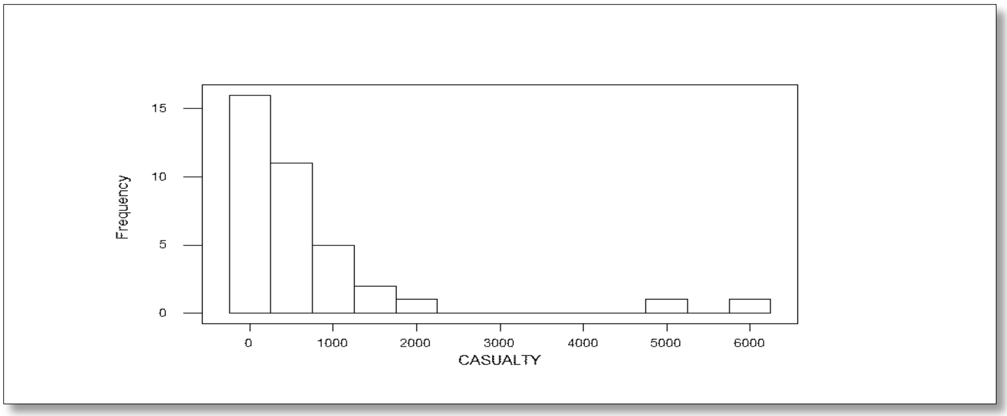

Figure 2. Histogram of the Number of Casualty

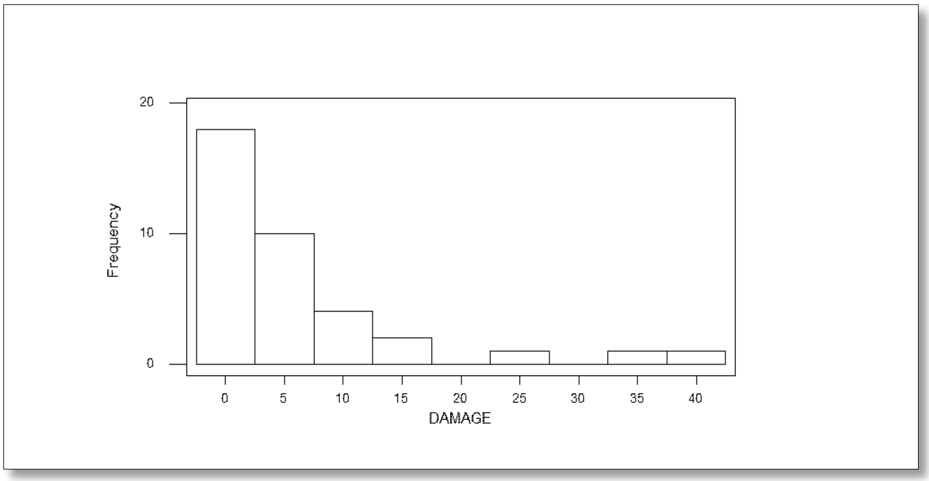

Figure 3. Histogram of the Estimated Damage

Both histograms reveal an exponential distribution with rate parameters $\beta_1 = 0.00137$ and $\beta_2 = 0.1531$ respectively. The Ryan- Joiner tests confirmed the hypothesis of exponentiality in both instances.

Representing the estimated damage and the number of casualties by D and C respectively, then the exponential behavior of these two random variables implies the existence of two (2) corresponding fractal variables X and Y given by:

19. $\frac{X}{\theta_1} = \exp(D)\ and\ \frac{Y}{\theta_2} = \exp(C)$

These fractal random variables can be interpreted as follows: X represents the impact to the economy of a locality sustaining a damage D when hit by a typhoon; Y represents the number of people directly affected when the number of casualties C is observed for the locality hit by a typhoon. We observe that a unit damage D (=1) translates into X = 2.7 (approximately 3) units of impact to the local economy and that this impact is exponentially growing with the sustained damage to infrastructure and agricultural production. Similarly, one casualty (C=1) affects Y=2.7 (roughly, 3 individuals) directly. In the Philippine context with the current family size of 3, this estimate is particularly informative.

Figures 4 and 5 show the histograms of X (economic impact) and Y (directly affected individuals by observed casualties).

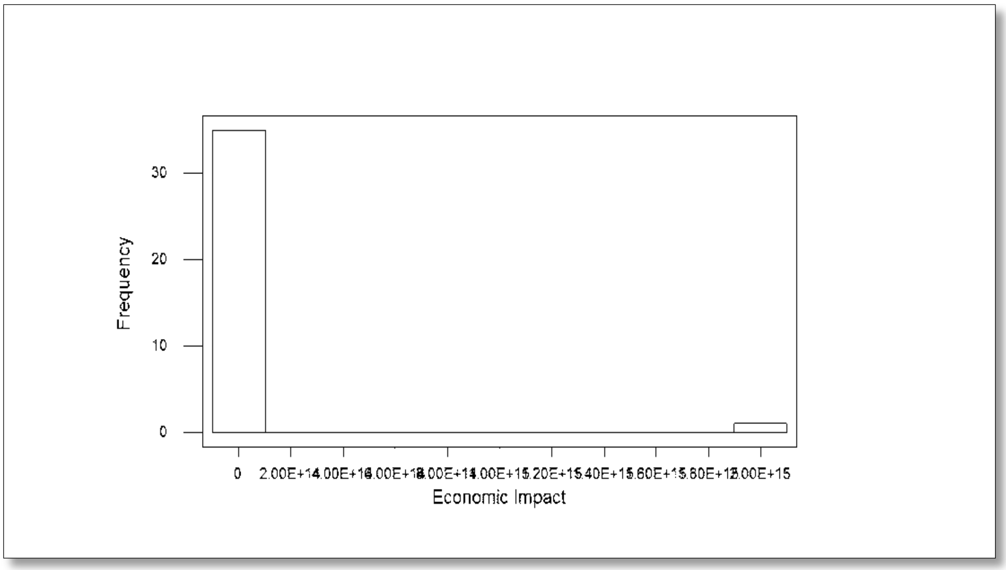

Figure 4. Histogram of the Economic Impact (X)

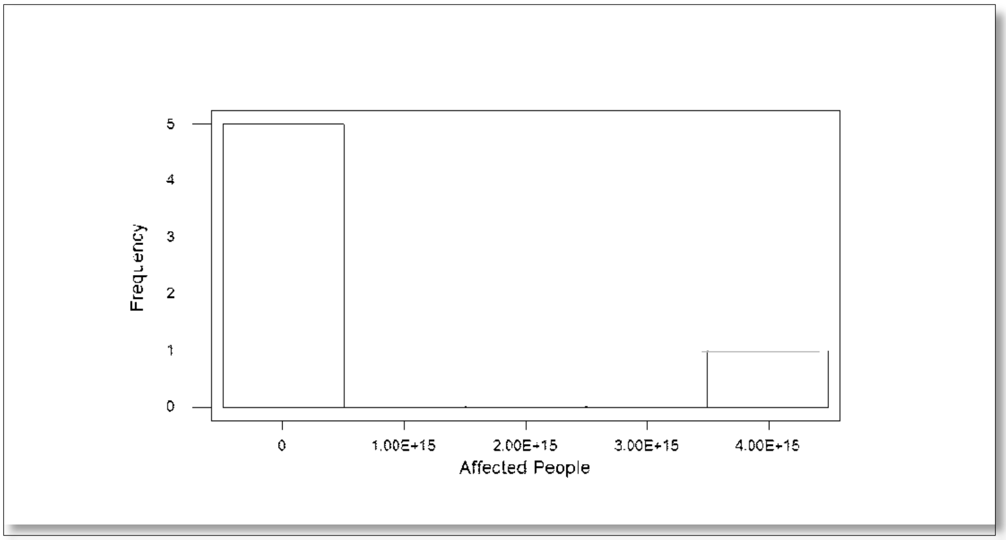

Figure 5. Histogram of the Number of Affected People (Y) by Casualties (C)

Of the two histograms, the first (Economic Impact) appears to behave more consistently with a power-law distribution (fractal distribution) than the second (Number of Affected People). The implication is that we are more certain about the extent of the economic impact of estimated damage (D) to infrastructure and agricultural production than about the number of people directly affected by casualties (C) of typhoon-led disasters.

Meanwhile, we computed the correlation between the random variables D and C, representing the correlation between two exponential random variates following the theory developed in Section 3. The correlation coefficient obtained was r (C,D) = 0.435 with a p value of p= 0.007. Thus, damage (D) and the number of casualties (C) are significantly correlated beyond the .01 probability level. Using the h (.) transfer function obtained in Section 3, we estimated the correlation of the fractal random variables X (economic impact) and Y (number of people affected) to be:

(20) $r(X,Y) = .584 + .413 \log(r(C,D)) = .584 + .413 \log(.435) = 0.2402.$

The result (20) can be empirically compared with the value of r (X,Y) from the data which is r (X,Y) = 0.2258 with a p-value of p = 0.625. The empirical formula for r (X,Y) and the data-based computed correlation are in close agreement. However, we note that since the fractal dimensions of X and Y are respectively 1.1531 and 1.00137 (both less than 3), these Pearson correlation estimates need to be viewed as a very rough estimate of the relationship between these two quantities.

In summary, we observed that while the estimated damage (D) and the number of casualties observed in a locality hit by typhoons correlate significantly, the corresponding economic impact (X) and numbers of people affected by the casualties (Y) are not. In order to link the latter two variables, it is necessary to obtain information about the casualties (C): whether or not they were primary breadwinners or not.

References

Ali, M., N. Mikhail and M. Haq (1978). *A Class of Bivariate Distributions Including the Bivariate Logistic*. (Journal of Multivariate Analysis, Vol. 8, pp. 405-412)

Iyer,S., D. Manjunath and R. Manivasakan, (2002). *Bivariate Exponential Distributions with Linear Structures*. (Sankhya: The Indian Journal of Statistics, Vol. 64, pp.156-166).

Kumar, P. (2010). *Probability Distribution and Estimation of the Ali-Mikhail-Haq Copula* (Applied Mathematical Sciences, Vol. 4, No. 14, pp. 657-666).

Lapinig, V. and J. Bolante (2014). *Typhoons, Casualty and Damage*. (Northwestern Mindanao State College Research Journal, CHED-JAS (candidate status), Vol. 2).

Marshall, A. and I. Olkin (1967). *A Generalized Bivariate Exponential Distribution*. (Journal of Applied Probability, Vol. 4, pp. 291-302)

Padua, R. (2014). *Fractal Statistical Analysis*. (Journal of Multidisciplinary Research, Surigao del Sur State University, CHED-JAS-B, Vol. 3, No. 1, pp. 98-107)