Mathematica Balkanica

Mathematical Society of South-Eastern Europe
A quarterly published by
the Bulgarian Academy of Sciences – National Committee for Mathematics

Mathematica
Balkanica

New Series Vol. 14, 2000, Fasc. 1-2

Statistical Studies on Fractal Dimensions

Sonoko Ishimura [1], *Sadao Ishimura* [2],
Yuichi Tamatsu [3], *Yoshinobu Ide* [3]

Presented by P. Kenderov

The study of fractal dimensions is one of the most useful methods for expressing the complexity of patterns. In this note, the fractal dimensions of the sagittal sutures of human craniums are analyzed by statistics. Their multiple regression equation is given and the most effective factors for their prediction are determined. Furthermore, some of the differences between the sagittal sutures of Caucasoids and of Mongoloids become clear.

AMS Subj. Classification: 28A80, 62H15, 62P25, 62P10

Key Words: fractal dimension, complexity of patterns, multiple regression analysis, human craniums

1. Fractal Dimension

The study of fractal dimensions is one of the most useful methods to express the complexity of not only mathematical patterns but also natural patterns. There are several definitions of fractal dimensions, but, for any fractal dimension, $1 = dim_F(l) \leq dimF(l) \leq 2$ holes where $dim_F(l)$ is a fractal dimension of a plane curve l and $dimT(l)$ is a topological dimension of l. A plane curve l with a large fractal dimension has a more complex pattern. In this note, we define a fractal dimension (also called the divider dimension) for a plane curve l as follows.

Let $N(d)$ be the number of steps required by a pair of dividers set at length d to traverse curve l. If there are two constants c and k such that $N(d) = cd^{-k}$ for any d, a fractal dimension of l is defined and we denote $k = dimF(l)$.

We can apply this idea to natural curves. A fractal dimension of a natural curve I can be estimated as the gradient of a log-log graph plotted over a suitable range of d and we define that a natural curve is fractal if the coefficient of determination of the regression line of its log-log graph is more than 0.90 [3].

We have collected the data from twenty sagittal sutures of human craniums (Table 1). Ten of them are typically Mongoloid (group M) and the other ten are typically Caucasoid (group C). For each sagittal suture curve, the number of steps $N(d)$ ($d = 3,5,7,9,11$) was measured.

Table 1: Measurements of sutura sagittalis

Cranium	3 mm	5 mm	7 mm	9 mm	11mm	dim_F
M_1	48	27	18	14	11	1.13
M_2	53	28	20	14	12	1.16
M_3	45	25	18	13	10	1.14
M_4	53	31	22	14	11	1.22
M_5	77	46	26	19	13	1.37
M_6	52	29	19	15	12	1.13
M_7	56	32	19	14	12	1.23
M_8	66	31	22	15	11	1.35
M_9	69	35	24	18	12	1.30
M_{10}	60	34	19	15	12	1.27
C_1	69	36	21	15	11	1.42
C_2	72	35	24	16	10	1.46
C_3	81	45	22	15	12	1.54
C_4	52	28	18	13	11	1.22
C_5	64	34	22	14	10	1.42
C_6	84	38	26	17	12	1.47
C_7	66	35	21	14	11	1.40
C_8	108	50	30	18	15	1.56
C_9	66	31	21	13	10	1.45
C_{10}	83	38	26	16	12	1.48

2. Regression line and fractal dimension

A regression line of plots is determined in order to minimize the sum of squares of the distance between a point and a line. The regression lines of log-log plots of our data are calculated and subsequently the coefficients of determination. The latter are near 1 and this shows the regression lines are a good fit. So all the lines of the sagittal sutures in our data are fractal and each fractal dimension is given as an absolute value of a gradient of the regression line (Table 1).

3. Prediction of fractal dimension

3.1. Multiple regression line

In this paragraph, fractal dimensions of sagittal sutures for each group are predicted by multiple regression lines. We set a fractal dimension as a dependent variable y and we denote the independent variables x_1, x_2, x_3, x_4 and x_5, where x_i $(i = 1, 2, 3, 4, 5)$ is the value $N(d)$ for d $(d = 3,5,7,9,11$ mm), respectively. In a regression line

$$y = b_1x_1 + b_2x_2 + b_3x_3 + b_4x_4 + b_5x_5 + b_0$$

the coefficients b_i $(i = 1, 2, 3, 4, 5)$ are called partial regression coefficients and b_0 is a constant. These coefficients are determined to minimize the sum of squares of residual between predicted values and observed values. Our result is shown in Table 2. We get the following multiple regression equations y_M, y_C and y of fractal dimensions for group M, C and all, respectively,

$$\begin{aligned} y_M = 1.342 &\times 10^{-2}x_1 + 4.882 \times 10^{-3}x_2 - 3.903 \times 10^{-3}x_3 \\ &-2.731 \times 10^{-2}x_4 - 3.561 \times 10^{-2}x_5 + 1.204, \end{aligned}$$

$$\begin{aligned} y_C = 1.183 &\times 10^{-2}x_1 + 5.882 \times 10^{-3}x_2 - 8.035 \times 10^{-3}x_3 \\ &-1.848 \times 10^{-2}x_4 - 6.614 \times 10^{-2}x_5 + 1.562, \end{aligned}$$

$$\begin{aligned} y = 1.066 &\times 10^{-2}x_1 + 7.831 \times 10^{-3}x_2 - 7.687 \times 10^{-3}x_3 \\ &-1.243 \times 10^{-2}x_4 - 6.321 \times 10^{-2}x_5 + 1.444. \end{aligned}$$

Their coefficients of determination are 0.994, 0.994 and 0.987. This shows the multiple regression equations fit the data well.

3.2. Test of the multiple regression equation

The multiple regression equations which are given are tested using analysis of variance. The hypothesis is that there is no linear relationship between fractal dimensions and the variables x_i $(i = 1, 2, 3, 4, 5)$. According to this test, the hypothesis for each case is rejected $(p = 0.05)$. So the fractal dimensions of sagittal sutures can be predicted with these multiple regression equations.

3.3. Test of partial regression coefficients

Each standardized partial regression coefficient bi (i = 1,2,3,4,5) expresses the effective level of its independent variable x_i to a dependent variable y. For group M, b_1 has the biggest absolute value. For group C, b_1 and b_5 have the biggest absolute values. This shows that the measurement at scale 3 mm is the most important for group M and the ones at scale 3mm and 11 mm are the most important for group C. These lengths denote the characteristics of waves for each group of sagittal sutures.

Table 2: Coefficients of multiple regression lines

Group		Unstandardized Coeff-s	Standardized Coeff-s	Significance	R^2
M	b_0	1.204		0.000	0.994
	b_1	1.342×10^{-2}	1.504	0.000	
	b_2	4.882×10^{-3}	0.319	0.039	
	b_3	-3.903×10^{-3}	−0.118	0.271	
	b_4	-2.731×10^{-2}	−0.583	0.005	
	b_5	-3.561×10^{-2}	−0.335	0.007	
C	b_0	1.562		0.000	0.994
	b_1	1.183×10^{-2}	1.978	0.003	
	b_2	5.882×10^{-3}	0.410	0.055	
	b_3	-8.035×10^{-3}	−0.301	0.230	
	b_4	-1.848×10^{-2}	−0.335	0.038	
	b_5	-6.614×10^{-2}	−1.083	0.000	
all	b_0	1.444		0.000	0.987
	b_1	1.066×10^{-2}	1.158	0.000	
	b_2	7.831×10^{-3}	0.366	0.001	
	b_3	-7.687×10^{-3}	−0.179	0.095	
	b_4	-1.243×10^{-2}	−0.155	0.034	
	b_5	-6.321×10^{-2}	-0.537	0.000	

Next, each partial regression coefficient is tested. For each variable x_i $(i = 1,2,3,4,5)$, the hypothesis $\beta_i = 0$ where $y = \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \beta_4 x_4 + \beta_5 x_5 + \beta_0$ is the model for the multiple regression equation. The result is shown in Table 2. In group M, the hypothesis is rejected for $\beta_1, \beta_2, \beta_4$ and β_5 at the level of significance 0.05, but not rejected for β_3. This means that the independent variables x_1, x_2, x_4 and x_5 contribute to the prediction for fractal dimensions, but x_3 does not contribute. In group C, the hypothesis is rejected for β_1, β_4 and $\beta_5 (p = 0.05)$, but not rejected for β_2 and β_3. This means that the independent variables x_1, x_4 and x_5 contribute to the prediction of the fractal dimensions, but x_2 and x_3 do not. These lengths also denote the characteristics of each group.

4. Analysis for the difference

4.1. Test of the difference

From our data of Table 1, the means of the two groups M and C are 1.230 and 1.442, respectively. The difference between these means is tested. The hypothesis is that $\mu_M = \mu_C$ where μ_M and μ_C are the population means of group M and C, respectively. The hypothesis is rejected and $\mu_M \neq \mu_C$ is determined. This means that there is a significant difference between the fractal dimensions of the two groups M and C.

Table 3: Coefficients of discriminant function

	Unstandardized Coefficients	Standardized Coefficients
a_1	-0.226	-2.940
a_2	-0.012	-0.076
a_3	0.219	0.679
a_4	0.649	1.163
a_5	1.202	1.245
a_6	-10.910	

4.2. Discriminant analysis

When discriminant analysis is used, the population variance-covariance ma-trices of both groups should be tested. This test is called Box's M-test. The hypothesis is that the population variance-covariance matrices of both groups are equal. If the hypothesis is not rejected, a linear discriminant function is used in the discriminant analysis. If the hypothesis is rejected, a Mahalanobis distance method is used. According to this test for our data, the hypothesis cannot be rejected and so a linear discriminant function is used.

Put

$$z = a_1x_1 + a_2x_2 + a_3x_3 + a_4x_4 + a_5x_5 + a_6,$$

the linear discriminant function between group M and C, where a_i $(i = 1, 2, 3, 4, 5, 6)$ is a constant. Each ai is determined to maximize the ratio S_B/S_T where S_B is the variation between group M and C, and S_T is the total variation. The result is shown in Table 3 and the following linear discriminant function is given:

$$z = -0.226x_1 - 0.12x_2 + -0.219x_3 + 0.649x_4 + 1.020x_5 - 10.910.$$

Using this function, 90% of craniums in our data are classified correctly. The coefficients of standardized linear discriminant function, a_1, a_4 and a_5 have bigger absolute values (Table 3). This means the measurements at scale 3 mm, 9 mm and 11 mm of the sagittal sutures are more important for the classification of the two groups.

4.3. Logistic regresson analysis

To calculate the probability that a cranium belongs to group C, logistic regression analysis is used instead of discriminant analysis.

Our data are not large enough to use all the variables x_1, x_2, x_3, x_4 and x_5. So we set a probability p as a dependent variable of the independent variables x_1 and x_5 which make a greater contribution to the discriminance. That is,

$$\log \frac{p}{1-p} = \beta_1 x_1 + \beta_5 x_5 + \beta_0,$$

or

$$p = \frac{\exp(\beta_1 x_1 + \beta_5 x_5 + \beta_0)}{1 + (\beta_1 x_1 + \beta_5 x_5 + \beta_0)}.$$

The coefficients β_i, $(i = 0, 1, 5)$ are determined by using maximum-likelihood method. The result is shown on Table 4 and we get

$$\log \frac{p}{1-p} = 0.2472 x_1 - 2.4630 x_5 + 11.8671.$$

Table 4: Coefficients of logistic regression equation

Coefficients	B	Significance
β_1	0.2472	0.0117
β_5	2.4630	0.0373
β_0	11.8671	0.1735

If $0 < p < 0.5$ for a cranium, it belongs to group M. If $0.5 < p < 1$, it belongs to group C. Using this logistic equation, 90% of craniums in our data are classificated correctly. We should notice that this analysis is used for only two groups. We cannot use this analysis for discriminating more than two races.

5. Conclusion

In forensic antholopology, it is important to be able to recognise racial characteristics. From our statistical analysis, we have the following conclusions about the sagittal sutures of human craniums.

1. The means of the fractal dimensions in group M (Mongoloid) and group C (Caucasoid) are 1.23 and 1.44, respectively. There is a significant difference between these fractal dimensions. The sagittal suture line of Cau-casoid is more complex than one of Mongoloid.

2. Fractal dimensions of sagittal sutures of Mongoloid and Caucasoid crani-ums are estimated by the following multiple regression equations, respectively;

$$\begin{aligned} y_M = 1.342 \quad & \times 10^{-2} x_1 + 4.882 \times 10^{-3} x_2 - 3.903 \times 10^{-3} x_3 \\ & -2.731 \times 10^{-2} x_4 - 3.561 \times 10^{-2} x_5 + 1.204 \end{aligned}$$

$$\begin{aligned} y_C = 1.183 \quad & \times 10^{-2} x_1 + 5.882 \times 10^{-3} x_2 - 8.035 \times 10^{-3} x_3 \\ & -1.848 \times 10^{-2} x_4 - 6.614 \times 10^{-2} x_5 + 1.562 \end{aligned}$$

In the group M (Mongoloid), the independent variable x_1 (measurement 3 mm) makes the biggest contribution to the prediction of the fractal dimension.In the group C (Caucasoid), the independent variables x_1 and x_5 (measurements 3 mm and 11 mm) make the biggest contribution to the prediction of the fractal dimension.

3. Whether craniums are Mongoloid or Caucasoid can be decided by the linear discriminant function

$$z = -0.226x_1 - 0.12x_2 + -0.219x_3 + 0.649x_4 + 1.020x_5 - 10.910.$$

If $z < 0$, a cranium is probably Mongoloid and if $z > 0$, it is probably Caucasoid. Furthermore the independent variables x_1, x_4 and x_5 (measurements 3 mm, 9 mm and 11 mm) make the biggest contribution to the discrimination betweem Mongoloid and Caucasoid.

4. Whether craniums are Mongoloid or Caucasoid can be decided by the value p, such that

$$\log \frac{p}{1-p} = 0.2472x_1 - 2.4630x_5 + 11.8671.$$

If $0 < p < 0.5$, a cranium is probably Mongoloid and if $0.5 < p < 1$, it is probably Caucasoid. Both variables x_1 and x_5 (measurements 3 mm and 11 mm) contribute to this discrimination significantly.

References

[1] A. J. Dobson, *An Introduction to Generalized Linear Models*, Chapman and Hall (1990).

[2] K. Falconer, *Fractal Geometry*, John Wiley & Sons (1990).
[3] Sonoko Ishimura, *On the Fractal Dimensions of Plane Sets*, Rep. of Chiba, Inst. of Technology (1991), (In Japanese).
[4] *SPSS Base 8.0, User's Guide.*
[5] *SPSS Professional Statistics 7.5.*

[1] *Chiba Institute of Technology, JAPAN*
[2] *Tsurumi University, JAPAN*
[3] *Tokyo Dental College, JAPAN*

Received: 15.05.1998