# RELATIONSHIP BETWEEN MORISITA'S MODEL FOR ESTIMATING THE ENVIRONMENTAL DENSITY AND THE GENERALIZED EULERIAN NUMBERS

# K. G. JANARDAN

Department of Mathematics, Eastern Michigan University, Ypsilanti, MI 48197, U.S.A.

(Received April 14, 1986; revised August 31, 1987)

Abstract. The environmental density has been defined (Morisita (1971, *Statistical Ecology*, the Pennsylvania State University Press, 379–401)) as the value of a habitat expressing its unfavorableness for settling of an individual which has a strong mutual-repulsive influence to other individuals in an environment. Morisita studied mutual repulsive behavior of ant lions (Glenuroides japanicus) and provided a recurrence relation without an explicit solution for the probability distribution of individuals settling in each of two habitats in terms of the environmental densities and the numbers of individuals introduced. In this paper the recurrence relation is explicitly solved; certain interesting properties of the distribution are discussed including its relation to the generalized Eulerian numbers and the estimation of the parameters.

Key words and phrases: Attraction and repulsion models, environmental density, Eulerian numbers, mle.

# 1. Introduction

Morisita (1971) was interested in measuring the habitat values related to an environment by individual ant lions. Ant lions are the larval forms of insects of the family Myremeleontidae. They are strongly mutually repulsive in that they prefer sparsely settled areas to areas populated with other ant lions. This might be perhaps due to the reduced chances of catching ants in crowded areas. However, an ant lion prefers also to dig pits in fine grained sand rather than rough grained sand, perhaps because the ants will more likely fall into such pits.

After a series of experimental studies with ant lions, Morisita presented the concept of environmental density by which the degree of preference for habitats can be quantitatively measured. He found that ant lions have a strong tendency to prefer fine to rough sand for digging pits when the population density is low. This tendency gradually decreased with increasing density until almost same number of ant lions settled in both types of sand exhibiting the existence of repulsive interference among individuals.

Shigesada *et al.* (1979) presented a new interpretation to the idea of environmental density proposed by Morisita. After developing a generalized differential equation, they showed that the solution of their model also gave the same results as that of Morisita. In order to explain his results quantitatively, Morisita assumed that the probability of settling in fine sand was proportional to the degree of unfavorableness of that habitat. In particular, Morisita used an experimental box containing fine sand in one half, side A, and coarse sand in the other half, side B. The ant lions were placed at the center of the box and allowed to settle where they chose. When a total of n ant lions had settled by digging a small pit somewhere in the box, the numbers in the fine sand (x) and the coarse sand (n-x) were counted. The experiment was repeated several times with n = 1, 2, ..., 7.

Then letting  $P_x(n)$  denote the probability that x of the n ant lions settling in fine sand, Morisita presented a recurrence relation (2.6). He did not get an explicit solution for it. However, he used it for the parameter estimation and for computing the probabilities numerically. In this paper, we obtain an explicit solution for  $P_x(n)$ ; establish certain interesting properties of the distribution, such as its relation to other discrete distributions, its relation to the generalized Eulerian numbers and moments recursion formula. The parameters of the model are estimated by both the methods of moments, and the maximum likelihood.

## 2. Morisita's model

Morisita modelled the experiment by supposing that the first ant lion in the box chose coarse sand with probability P(A) and fine sand with probability P(B) = 1 - P(A). These probabilities were defined as

(2.1) 
$$P(A) = a/(a+b), \quad P(B) = b/(a+b),$$

where the real numbers a and b are the degrees of unfavorableness or the environmental densities of the fine sand and coarse sand, respectively. Subsequent ant lions would most likely face a choice between crowded fine sand or sparsely settled coarse sand; Morisita postulated that

(2.2) Pr (*n*-th ant lion chooses coarse sand | x out of (n - 1)are already in fine sand) = (a + x)/(a + b + n - 1).

Thus, if any ant lions are in fine sand the probability of choosing coarse sand is increased. Similarly,

(2.3) Pr (*n*-th ant lion chooses fine sand 
$$|n - x - 1|$$
 are in coarse sand)  
=  $(b + n - x - 1)/(a + b + n - 1)$ .

Let  $\{L_n = x\}$  denote the event that x ant lions out of n settled in fine sand. Then

$$(2.4) P_x(n) = \Pr(L_n = x).$$

A recurrence relation can be found as follows. Define the indicator random variables,  $I_n$ ,

$$I_n = \begin{cases} 1 & \text{if the } n\text{-th ant lion chooses fine sand with probability} \\ (b+n-1-L_{n-1})/(a+b+n-1), \\ 0 & \text{otherwise}. \end{cases}$$

Then clearly,

(2.5) 
$$L_n = I_1 + I_2 + \dots + I_n = L_{n-1} + I_n$$

Now,

$$\{L_n = x\} = \{L_{n-1} = x, I_n = 0\} \cup \{L_{n-1} = x - 1, I_n = 1\},$$
  

$$\Pr\{L_n = x\} = \Pr\{L_{n-1} = x, I_n = 0\} + \Pr\{L_{n-1} = x - 1, I_n = 1\}$$
  

$$= \Pr\{I_n = 0 | L_{n-1} = x\} \Pr\{L_{n-1} = x\}$$
  

$$+ \Pr\{I_n = 1 | L_{n-1} = x - 1\} \Pr\{L_{n-1} = x - 1\}.$$

Thus,

(2.6) 
$$P_x(n) = \frac{a+x}{a+b+n-1} P_x(n-1) + \frac{b+n-x}{a+b+n-1} P_{x-1}(n-1),$$

with initial values

(2.7) 
$$P_0(1) = a/(a+b), \quad P_1(1) = b/(a+b).$$

We shall show in the next section that the solution of (2.6) with the initial conditions as given in (2.7) is

(2.8) 
$$P_x(n) = E_{n,x}(a,b)/(a+b)^{[n]}, \quad x=0,1,\ldots,n,$$

where  $E_{n,x}(a, b)$  are the generalized Eulerian numbers defined in the next section.

## 3. Generalized Eulerian numbers

DEFINITION. For given positive real numbers a and b, we define  $E_{n,x}(a,b)$  for all  $x \ge 0$  and  $n \ge 0$  by the formula

(3.1) 
$$E_{n,x}(a,b) = \sum_{j=0}^{x} (-1)^{j} c(x,j) (a+b)^{[x-j]} (a+b+n)^{(j)} (a+x-j)^{n} / x!$$

If a = 0 and b = 1, then  $E_{n,x}(a, b)$  becomes  $A_{n,x}$  where  $A_{n,x}$  are the wellknown Eulerian numbers introduced in 1736 by Euler (see Takacs (1979) for historical notes and the references contained therein):

(3.2) 
$$A_{n,x} = \sum_{j=0}^{x} (-1)^{j} c(n+1,j) (x-j)^{n}.$$

Therefore, we shall call  $E_{n,x}(a, b)$  as the generalized Eulerian numbers.

In formula (3.1) the extended factorials are defined as follows:

(3.3) 
$$m^{[r]} = m(m+1)(m+2)\cdots(m+r-1), \quad m^{[0]} = 1,$$
  
 $c(x,j) = x!/j!(x-j)!.$ 

THEOREM 3.1. The generalized Eulerian numbers  $E_{n,x}(a,b)$  as defined in (3.1) are obtained as the unique solution of the recurrence formula:

$$(3.4) E_{n,x}(a,b) = (a+x)E_{n-1,x}(a,b) + (b+n-x)E_{n-1,x-1}(a,b),$$

with the boundary conditions  $E_{1,0}(a,b) = a$ ,  $E_{1,1}(a,b) = b$ ,  $E_{n,0}(a,b) = a^n$ and  $E_{n,n}(a,b) = b^n$ .

**PROOF.** First, we shall show that the numbers  $E_{n,x}(a, b)$ , as defined in (3.1), satisfy the recurrence relation (3.4). On the right hand side of (3.1) substitute  $(a + x)(a + x - j)^{n-1} - j(a + x - j)^{n-1}$  for  $(a + x - j)^n$ , we get

(3.5) 
$$E_{n,x}(a,b) = (a+x) \sum_{j=0}^{x} (-1)^{j} c(x,j)(a+b)^{[x-j]} \cdot (a+b+n)^{(j)} (a+x-j)^{n-1} / x! - \sum_{j=0}^{x} (-1)^{j} c(x,j)(a+b)^{[x-j]}$$

• 
$$(a + b + n)^{(j)} j (a + x - j)^{n-1} / x!$$

If we substitute  $(a + b + n)^{(j)} = (a + b + n - 1)^{(j)} + j(a + b + n - 1)^{(j-1)}$  in the first term of (3.5) and  $(a + b + n)^{(j)} = (a + b + n)(a + b + n - 1)^{(j-1)}$  in the second term of (3.5), we obtain

$$(3.6) \quad E_{n,x}(a,b) = (a+x) \sum_{j=0}^{x} (-1)^{j} c(x,j)(a+b)^{[x-j]} \cdot (a+b+n-1)^{(j)}(a+x-j)^{n-1}/x! + (a+x) \sum_{j=0}^{x} (-1)^{j} c(x,j)(a+b)^{[x-j]} \cdot j(a+b+n-1)^{(j-1)}(a+x-j)^{n-1}/x! - (a+b+n) \sum_{j=0}^{x} (-1)^{j} c(x,j)(a+b)^{[x-j]} \cdot j(a+b+n-1)^{(j-1)}(a+x-j)^{n-1}/x! .$$

Combining the last two terms of (3.6), we get

$$(3.7) Ext{ } E_{n,x}(a,b) = (a+x) \sum_{j=0}^{x} (-1)^{j} c(x,j)(a+b)^{[x-j]} \\ \cdot (a+b+n-1)^{(j)}(a+x-j)^{n-1}/x! \\ + (b+n-x) \sum_{j=0}^{x-1} (-1)^{j} c(x-1,j)(a+b)^{[x-1-j]} \\ \cdot (a+b+n-1)^{(j-1)} \\ \cdot (a+x-1-j)^{n-1}/(x-1)! .$$

The first sum on the right hand side of (3.7) is  $E_{n-1,x}(a, b)$  and the second sum is  $E_{n-1,x-1}(a, b)$ . Thus, if  $E_{n,x}(a, b)$  is as defined in (3.1), the recurrence (3.4) is satisfied for  $0 \le x \le n$ .

It is easy to verify directly that the right hand side of (3.1) yields the values a and b for  $E_{1,0}(a, b)$  and  $E_{1,1}(a, b)$ , respectively. Since these initial boundary values are satisfied, it is easy to prove that (3.1) is the unique solution of (3.4) by showing that it satisfies (3.4) for x = 0, 1, 2, ..., n, and n = 2, 3, 4, ...

THEOREM 3.2. For given positive real numbers a and b, define  $E_{n,x}(a,b)$  for  $x \ge 0$  and  $n \ge 0$  by the recurrence formula (3.4) with the boundary conditions  $E_{n,0}(a,b) = a^n$  and  $E_{n,n}(a,b) = b^n$ . Then

K. G. JANARDAN

(3.8) 
$$\sum_{x=0}^{n} E_{n,x}(a,b) = (a+b)^{[n]}.$$

**PROOF.** For the sake of simplicity let us write  $E_{n,x}(a,b) \equiv E_{n,x}$  in (3.4). Then, we can rewrite (3.4) as

(3.9) 
$$E_{n+1,x} = (a+x)E_{n,x} + (b+n+1-x)E_{n,x-1}.$$

Summing both sides of (3.9) from 1 to n, w.r.t. x, we get

(3.10) LHS = 
$$\sum_{x=1}^{n} E_{n+1,x} = \sum_{x=0}^{n+1} E_{n+1,x} - E_{n+1,0} - E_{n+1,n+1}$$
  
=  $\sum_{x=0}^{n+1} E_{n+1,x} - a^{n+1} - b^{n+1}$ ,

(3.11) 
$$\mathbf{RHS} = \sum_{x=1}^{n} (a+x)E_{n,x} + \sum_{x=1}^{n} (b+n+1-x)E_{n,x-1}$$
$$= a\sum_{x=0}^{n} E_{n,x} - aE_{n,0} + \sum_{x=0}^{n} xE_{n,x} + (b+n+1)\sum_{y=0}^{n} E_{n,y}$$
$$- (b+n+1)E_{n,n} - \sum_{y=0}^{n-1} yE_{n,y} - \sum_{y=0}^{n-1} E_{n,y}$$
$$= (a+b+n)\sum_{x=0}^{n} E_{n,x} - aE_{n,0} - bE_{n,n}$$
$$= (a+b+n)\sum_{x=0}^{n} E_{n,x} - a^{n+1} - b^{n+1}.$$

Equating (3.10) and (3.11) we get

(3.12) 
$$\psi_{n+1} = (a+b+n)\psi_n$$
,

where  $\psi_n = \sum_{x=0}^n E_{n,x}$ . Solving (3.12) recursively, we can show that  $\psi_n = (a+b)(a+b+1)\cdots(a+b+n-1)\psi_0$ . Thus, since  $\psi_0 = 1$ ,  $\psi_n = \sum_{x=0}^n E_{n,x} = (a+b)^{[n]}$  as required.

Now dividing (3.4) by  $(a + b)^{[n]}$  and noting that  $(a + b)^{[n]} = (a + b + n - 1) \cdot (a + b)^{[n-1]}$ , we have

(3.13) 
$$\frac{E_{n,x}(a,b)}{(a+b)^{[n]}} = \frac{(a+x)}{a+b+n-1} \frac{E_{n-1,x}(a,b)}{(a+b)^{[n-1]}} + \frac{(b+n-x)}{a+b+n-1} \frac{E_{n-1,x-1}(a,b)}{(a+b)^{[n-1]}}.$$

Comparing (2.6) and (3.13) we get (2.8).

444

# 4. Relation with other discrete distributions

#### 4.1 A social attraction model

Instead of the (socially) mutually repulsive assumption built into the model through (2.2), we could develop a mutual attraction model as has been done by Morisita (1971), with the following postulate:

# (4.1) Pr (*n*-th ant lion chooses fine sand |x| out of n-1

are already in fine sand) = 
$$(a + x)/(a + b + n - 1)$$
,

where the positive real numbers a and b are indices of attractiveness of the habitats A and B. Then, corresponding to (2.6), we obtain the recurrence relation:

(4.2) 
$$P_x(n) = \frac{a+x}{a+b+n-1} P_{x-1}(n-1) + \frac{b+n-x-1}{a+b+n-1} P_x(n-1).$$

The solution of (4.2) is

(4.3) 
$$P_x(n) = c(n, x)(-a)^{(x)}(-b)^{(n-x)}/(-a-b)^{(n)}.$$

This probability distribution is known as the negative hypergeometric distribution, which is quite different from (2.8). This has a closed form single expression where as  $P_x(n)$  in (2.8) involves the sum of a finite number of terms.

The social repulsion and social attraction models can be related to the following urn sampling scheme.

# 4.2 Urn models

An urn contains "b" white and "a" black balls. A ball is drawn at random and then replaced together with an additional ball of the *opposite* color. The procedure is repeated n times. Let  $L_n$  denote the number of white balls observed in n draws, with  $P_x(n) = \Pr \{L_n = x\}$ .

The event  $\{L_n = x\}$  can occur in exactly one of two mutually exclusive ways. Either  $\{L_{n-1} = x\}$  has already occurred and a black ball is drawn at the *n*-th draw or  $\{L_{n-1} = x - 1\}$  has occurred and a white ball is drawn at the *n*-th draw. Thus, using standard rules of probability, we obtain a recurrence relation which is identical to the relation (2.6) obtained earlier for Morisita's social repulsion model.

The social attraction model can be derived starting with "a" white and "b" black balls initially in the above urn. This time n balls are selected at random such that at each draw the ball drawn is replaced together with an additional ball of the *same* color. Then,  $L_n$ , the number of white balls

observed in n trials, obeys the recurrence relation (4.2) and the probability distribution (4.3).

# 5. Moments of $P_x(n)$

Let  $\mu(n, r)$  denote the r-th factorial moment of the distribution. That is

(5.1) 
$$\mu(n,r) = \sum_{x=0}^{n} x^{(r)} P_x(n) .$$

From (2.6) we can find a recursion formula for the factorial moments (5.1),

(5.2) 
$$\mu(n,r) = \frac{a+b+n-r-1}{a+b+n-1} \mu(n-1,r) + \frac{(b+n-r)}{a+b+n-1} \mu(n-1,r-1)$$

Specializing it for r = 1 and r = 2 we can derive

(5.3) 
$$\mu(n,1) = (nb+c_2)/(a+b+n-1),$$

which is the same as the one obtained by Morisita for the expected value of settled number of individuals, and

(5.4) 
$$\mu(n,2) = [2b^2c_2 + (6b+2)c_3 + 6c_4]/(a+b+n-1)(a+b+n-2).$$

The variance,  $\sigma^2$ , is obtained from (5.3) and (5.4), using the fact that  $\sigma^2 = \mu(n, 2) + \mu(n, 1) - [\mu(n, 1)]^2$ , and is given by

(5.5) 
$$\sigma^{2} = \frac{nab(a+b-1)+c_{2}(a+b)^{2}+2c_{3}(a+b)+(n-1)c_{3}/2}{(a+b+n-1)^{2}(a+b+n-2)}$$

where  $c_j = c(n, j)$  for j = 2, 3, 4 as defined in (3.3).

## 6. Estimation of parameters a and b

For estimating the environmental densities a and b, Morisita (1971) utilized the ratio of the mean number of individuals in each of the two sands to the total number of individuals introduced. Here, we first utilize the method of moments for estimating the parameters a and b. Then we consider the maximum likelihood method.

## 6.1 Method of moments

Let m and  $s^2$  denote the mean and variance of the sample. Then equating these to the corresponding population moments, we obtain

(6.1) 
$$a = [nb + c_2]/m - b - n + 1,$$

where b is the positive root of the cubic equation,

(6.2) 
$$a_0b^3 + a_1b^2 + a_2b + a_3 = 0$$
.

Here,

(6.3)  

$$a_{0} = (n^{2}/m)(n/m-1) - n^{3}s,$$

$$a_{1} = 3n^{2}c_{2}/m^{2} - n(3n^{2}+c_{2})/m + 2n^{2}(1-2c_{2}s^{2}+s^{2}),$$

$$a_{2} = [3nc_{2}^{2} + 2nm(1-nc_{2})]/m^{2} - n(3nc_{2}-2c_{3})/m + n^{2}(2n-1) - nc_{2}s^{2}(3c_{2}-2),$$

$$a_{3} = c_{2}(c_{2} - nm + m)^{2}/m^{2} + 2c_{3}(c_{2} - nm + m)/m + (n-1)c_{3}/2 - (c_{2}-1)s^{2}c_{2}^{2}.$$

#### 6.2 Maximum likelihood method

Assuming that each ant lion responds independently, the likelihood function of the observed outcome (as in Table 2) is given by

(6.4) 
$$L(a, b, \mathbf{x}) = \prod_{n=1}^{k} \binom{M_n}{x_n} P_n^{x_n} (1 - P_n)^{M_n - x_n},$$

where  $M_n$  is the total number of individuals introduced into the experimental box with *n* individuals (n = 1, 2, ..., k, k = 7) in each experiment;  $x_n$ is the total number of individuals (out of  $M_n$ ) entering the fine sand, and  $P_n$ is the probability of success. Success is defined as observing at least one ant lion in fine sand. Thus,

(6.5)  $P_n = \Pr$  (at least one in fine sand |n| individuals are introduced)

$$=\prod_{x=1}^{n} P_{x}(n) = 1 - P_{0}(n) = 1 - a^{n}/(a+b)^{[n]}.$$

Substituting (6.5) in (6.4), the likelihood function becomes

(6.6) 
$$L(a, b, \mathbf{x}) = \prod_{n=1}^{k} \binom{M_n}{x_n} \left[ 1 - \frac{a^n}{(a+b)^{[n]}} \right]^{x_n} \left[ \frac{a^n}{(a+b)^{[n]}} \right]^{M_n - x_n}.$$

Maximization of equation (6.6) is not possible in a closed form since it is a non-linear function of a and b. Therefore, the parameters a and b can be obtained iteratively, using numerical optimization techniques.

## 6.3 Numerical example

The observed data in Tables 1 and 2 are from Tables 1 and 2 of Morisita (1971). For the observed data in Table 1, first, moment estimate of b is obtained (for each different type of experiment) as a root of the cubic (6.2) using Newton-Raphson iteration. The value of a is then obtained after substituting the value of b in (6.1). Final moment estimates of a and b are taken as the averages of the seven separate estimates of a and b to obtain:

(6.7) 
$$\tilde{a} = 0.2355, \quad \tilde{b} = 0.7647.$$

The over all fit of the model to the data (in Table 1) is encouraging as determined by the low chi-square values. Moreover, our method gives

Number introduced	Number of ant lions in fine sand									
n	$x \rightarrow$	0	1	2	3	4	5	6	7	
1	0	3	29							
	MM	7.53	24.47							
	MLE	7.55	24.45							
2	0	0	19	13						
	MM	0.89	21.76	9.35						
	MLE	0.91	21.58	9.51						
3	0	0	7	24	1					
	MM	0.07	9.78	19.77	2.38					
	MLE	0.07	9.68	19.77	2.48					
4	0	0	3	17	10	0				
	MM	0	2.89	16.69	9.98	0.44				
	MLE	0	2.86	16.57	10.10	0.46				
5	0	0	0	10	15	4	0			
	MM	0	0.69	9.32	15.17	3.76	0.06			
	MLE	0	0.68	9.22	15.17	3.85	0.06			
6	0	0	2	4	5	9	2	0		
	MM	0	0.11	3.05	10.64	7.31	0.88	0.00		
	MLE	0	0.11	3.02	10.58	7.38	0.91	0.001		
7	0	0	0	1	2	4	3	0	0	
	MM	0	0.01	0.48	3.18	4.61	1.61	0.10	0	
	MLE	0	0.01	0.48	3.15	4.15	1.64	0.11	0	

Table 1. Observed and expected distributions of ant lions by methods of moments and maximum likelihood.

O stands for observed frequency, MM denotes expected frequency by moments and MLE denotes expected frequency by maximum likelihood.

Ant lions		Number of ant lions introduced $m_n$	Number in fine sand				
in each	Number of experiments		Observed	ved Estimated			
n	enperments		Xn	Morisita's	MM	MML	
1	32	32	29	28.3	24.5	24.5	
2	32	64	45	42.6	40.5	40.6	
3	32	96	58	58.1	56.5	56.7	
4	30	120	67	69.3	68.0	68.1	
5	29	145	81	81.3	80.2	80.3	
6	22	132	71	72.6	71.8	72.0	
7	10	70	39	38.0	37.6	35.9	

Table 2. Total number of ant lions in fine sand for a total of 187 experiments.

more or less the same over all results as that of Morisita.

Next, we considered the maximum likelihood method. As stated earlier, the likelihood equation does not provide closed form solutions for a and b. We therefore, utilized Fletcher-Reeve (1964) optimization algorithm to obtain a global fit:

(6.8) 
$$\hat{a} = 0.2449 \pm (0.315), \quad \hat{b} = 0.7928 \pm (0.422),$$

with an estimated asymptotic correlation of 0.943. The values in the parenthesis are asymptotic standard errors (obtained using formulas in Kendall and Stuart ((1977), pp. 246-248) of the estimates of a and b, respectively. We can see the advantages of this approach over the method of moments. The fitted values obtained by using estimates of a and b in (6.8) are given in Table 1.

In conclusion, it may be noted that the maximum likelihood estimates are very close to our moment estimates. However, although our estimate of b (by both methods) is of the same order as that of Morisita's (his estimate of b is 0.644), our estimate of a is about three times larger than that of his estimate (a = 0.086). As a consequence, our estimate of the degree of unfavorableness to fine sand is much higher than that of Morisita and hence our estimates of the number of ant lions (Table 2) in fine sand are a bit lower than those of Morisita.

#### Acknowledgement

The author wishes to thank the referee for helpful, constructive comments and suggestions which have helped improve the revision.

#### K. G. JANARDAN

#### REFERENCES

- Fletcher, R. and Reeve, C. M. (1964). Function minimzation by conjugate gradients, *Comput. J.*, 7, 149-154.
- Kendall, M. G. and Stuart, A. (1977). The Advanced Theory of Statistics, Vol. 1, 4th ed., MacMillan Publishing Company, New York.
- Morisita, M. (1971). Measuring of habitat value by environmental density method, *Statistical Ecology*, (eds. G. P. Patil *et al.*), 379-401, The Pennsylvania State University Press.
- Shigesada, N., Kawasaki, K. and Teramato, E. I. (1979). Spatial segregation of interacting species, J. Theoret. Biol., 79, 83-99.
- Takacs, L. (1979). A generalization of the Eulerian numbers, Publ. Math. Debrecen, 26, 173-181.