J. mar. biol. Ass. India, **46** (1) : 120 - 123, Jan. - June, 2004

NOTE

Fitting allometric model using fuzzy least squares

V.M. Kandala and Prajneshu

Indian Agricultural Statistics Research Institute, New Delhi – 110 012, India

Abstract

Allometric model, in logarithm form, is generally fitted using "Method of least squares". This methodology is applicable only when response variable, viz. weight of a fish is assumed to be described by a single fixed number corresponding to each fixed value of explanatory variable, viz. length of a fish. But in reality, due to fuzziness or vagueness of underlying phenomenon, weights of fish of exactly equal lengths need not be same. Thus, response variable is fuzzy or vague as it can be described only by an interval. To deal with this situation, "Method of fuzzy least squares" is discussed. Finally, the methodology is illustrated on some length-weight data of fish culled from literature.

Allometric model is extensively employed for determining length-weight relationship that serves two-fold purpose: (i) converting length measurements to weight estimates, and (ii) computing condition indices regarding well-being of fish. In reality, weights of fish of even equal lengths are not exactly same, i.e. these vary and so can be described only by an interval. As no methodology was available until now to deal with a situation in which data pertaining to response variable is vague or fuzzy, the only way out was to somehow restrict the data to precise or crisp form. However, by doing so, a lot of vital information about the underlying phenomenon is undoubtedly lost. Fortunately, a new technique called "Fuzzy regression methodology (FRM)" has recently been developed (Kacprzyk and Fedrizzi, 1992) which is capable of handling fuzzy data. The objective of the present paper is to discuss this approach and apply it to some real data so that fishery scientists may start using it for determining length-weight relationship more efficiently.

Material and methods

The well-known allometric model is usually expressed, in logarithmic form, as

$$W = a + b L + \in \qquad \ldots(1)$$

Here L and W represent respectively logarithms of length and weight of a fish and $\in$ is the error term. Parameters 'a' and 'b' are estimated employing "Method of least squares" (Draper and Smith, 1998). In eq.(1), it is assumed that response variable W can be expressed by a single number for every fixed value of explanatory variable L. However, in reality, for every fixed value of L, the variable W lies in an interval. In FRM (Ishibuchi, 1992), first step is to replace eq.(1) by following equation :

$$< W_c, W_r > = < a_c , a_r > + < b_c , b_r > L \quad ...(2)$$

The essential difference is that now the parameters 'a' and 'b' and consequently 'W' are each expressed by an interval and so each one of these is described by two quantities; subscripts 'c' and 'r' indicate respectively centre and radius or width. Thus, for example, parameter 'b' is now fuzzy and is expressed by a centre value b_c and radius b_r. It may be noted that, in eq.(2), the explanatory variable 'L' is crisp and not fuzzy.

In order to estimate parameters of eq.(2), "Method of fuzzy least squares" (Diamond, 1988) is employed. In this approach, parameters are estimated by minimizing total vagueness, i.e. sum of widths of predicted intervals under the constraints that all data points fall within estimated interval of response variable.

Eq.(2) can be expressed as (Savic and Pedrycz, 1991) :

$$W_c = a_c + b_c . L \quad ...(3)$$

and

$$W_r = a_r + b_r . | L | \quad ...(4)$$

Now total vagueness in the model with m data points is $\sum_{j=1}^{m} W_{rj}$. The problem of estimation can be viewed as a "Linear programming problem" (LPP) (Taha, 1997) and can be formulated as follows:

Minimize

$$\sum_{j=1}^{m} \left[a_r + b_r \, | L_j | \right]$$

Subject to

$$\sum_{j=1}^{m} \left[\left(a_c + b_c . L_j \right) - \left(a_r + b_r . | L_j | \right) \right] \leq W_j$$

$$\sum_{j=1}^{m} \left[\left(a_c + b_c . L_j \right) + \left(a_r + b_r . | L_j | \right) \right] \geq W_j$$

and

$$a_r , b_r \geq 0$$

For crisp data, least square estimators satisfy optimum properties. Therefore, it is logical to set centre values in fuzzy least squares (FLS) method equal to corresponding least squares (LS) estimates. Accordingly, estimates of a_c and b_c are taken respectively as LS estimates of parameters a and b appearing in eq.(1). These can be obtained by employing any standard statistical software package, such as Statistical Analysis System. Now the remaining two quantities, viz. a_r and b_r are to be estimated. To this end, "Simplex procedure" (Taha, 1997) is generally employed for solving the above LPP. Many software packages are available for achieving the task. Some of these, such as LP88, LINDO are specially meant for solving LPP, whereas general purpose Statistical Analysis System (SAS, 1998) software package has an inbuilt procedure, viz. Proc LP. However, any standard spreadsheet package, like Microsoft Excel can also be used to solve LPP manually.

Results and discussion

To illustrate above methodology, length-weight data of fish given in Jayaprakash (1999) is considered and is reproduced in Table 1 for ready reference. Using "Method of least squares" to above data, eq.(1) yields

$$W = -11.99 + 2.96 \ L \quad \ldots(6)$$
$$\quad\quad (0.38) \quad (0.08)$$

Figures in parentheses represent corresponding standard errors of estimates. Now putting $a_c = -11.99$ and $b_c = 2.96$ in eq.(5), the LPP is

Minimize

$$[a_r + b_r \ (4.38 + 4.44 + \ldots + 5.04)\ ]$$

Subject to the following sets of conditions

$(-11.99 + 2.96 \times 4.38) - (a_r + 4.38\ b_r) \leq 1.12$

(and so on)

$(-11.99 + 2.96 \times 5.04) - (a_r + 5.04\ b_r) \leq 3.04$

and

$(-11.99 + 2.96 \times 4.38) + (a_r + 4.38\ b_r) \geq 1.12$

(and so on)

$(-11.99 + 2.96 \times 5.04) + (a_r + 5.04\ b_r) \geq 3.04$

and

$a_r\ ,\ b_r \geq 0$

Employing SAS (1998) software package to solve the above LPP, the estimated model is

$$< W_c\ ,\ W_r > \ = \ < -11.99,\ 0.14 > + < 2.96\ ,\ 0 > L \quad \ldots(7)$$

The estimated weights corresponding to various observations on lengths, computed using eqs.(6) or (7), are reported in the third column of Table 1. Further, predicted intervals for LS and FLS estimates are computed using the following formulae :

$$(4.24 \times 10^{-6}\ L^{2.88}, \quad 9.07 \times 10^{-6}\ L^{3.04}\ )$$

and

$$(\ 5.40 \times 10^{-6}\ L^{2.96}, \quad 7.14 \times 10^{-6}\ L^{2.96}\ )$$

Widths of these are given respectively in fourth and fifth columns of Table 1. A

Table 1. *Fitting of fuzzy allometric model*

Length (mm)	Weight (g)	Estimated weight (g)	Width of predicted interval (g.)	
			Least squares	Fuzz least squares
80	3.05	2.68	4.32	0.76
85	3.07	3.20	5.21	0.91
90	3.68	3.79	6.22	1.08
95	4.56	4.45	7.35	1.27
100	4.72	5.18	8.61	1.47
105	6.10	5.99	10.01	1.70
110	6.65	6.87	11.56	1.95
115	7.65	7.84	13.26	2.23
120	9.16	8.89	15.12	2.53
125	10.14	10.03	17.15	2.85
130	10.43	11.27	19.35	3.20
135	12.99	12.60	21.74	3.58
140	14.48	14.03	24.32	3.99
145	15.32	15.57	27.10	4.43
150	17.35	17.21	30.09	4.89
155	20.90	18.96	33.29	5.39

perusal indicates that average widths of predicted intervals for LS and FLS estimates are respectively 15.92 and 2.64. To get visual idea, predicted intervals for LS and FLS procedures along with data, are exhibited in Figure 1.

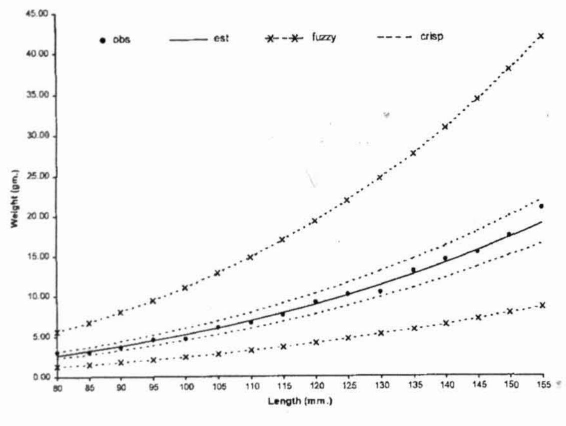

Fig. 1. *Predicted intervals for least squares (—) and fuzzy least squares (—) procedures along with data (♦)*

Similar results are also obtained when length-weight data of ovary of fish, as reported in Jayaprakash (1999), are considered. Again, average widths of predicted intervals in respect of FLS procedure is almost one-fifth of that provided by LS procedure. However, details are omitted here to save space.

Thus, predicted intervals computed using "Method of fuzzy least squares" have much shorter average widths as compared to that obtained using "Method of least squares". This implies that former procedure is more efficient than latter. However, this aspect is only of secondary importance. The main message emerging out of present paper is that correct methodology to determine length-weight relationship in fish is that of "Fuzzy least squares" rather than ordinary "Least squares".

References

Diamond, P. 1988.. *Sci.*, **46** : 141-157.

Draper, N.R. and H.Smith. 1998. Applied regression analysis, 3rd Edn., John Wiley, New York. p. 217-34

Ishibuchi, H. 1992. *Japan J. Fuzzy Theory & Systems*, **4** : 137-148.

Jayaprakash, A. A. 1999. *J. mar. biol. Ass. India*, **41** : 88-92.

Kacprzyk, J. and M. Fedrizzi. 1992. Fuzzy regression analysis, Omnitech Press, Warsaw.279 pp.

SAS 1998. SAS/OR user's guide, SAS Institute, North Carolina.p.234-39.

Savic, D. and W. Pedrycz. 1991. *Fuzzy Sets and Systems*, **39** : 51-63.

Taha, H.A. 1997. Operations research: An introduction, 6th Edn., Prentice Hall of India, New Delhi. p. 67-109.