SOME EXAMPLES OF DOUBLE EXPONENTIAL FITNESS FUNCTIONS

B. F. J. MANLY

Biometrics Unit, University of Otago, P.O. Box 56, Dunedin, New Zealand

SUMMARY

The advantages of using a double exponential fitness function are pointed out for the analysis of differential survival data. The uses of this type of function are illustrated on Bumpus' (1898) data on the survival of English sparrows after a severe storm and also on Cook and O'Donald's (1971) data on the over-winter survival of a large sample of snails.

1. INTRODUCTION

A FITNESS function gives the probability of an individual surviving a period of natural selection according to the values that it has for one or more characters $X_1, X_2, ..., X_m$. It will seldom, if ever, be possible to know the precise mathematical form for this function and in this note it is suggested that it is convenient to assume a double exponential type of function when the data available for estimation consists of information on which of n individuals survived one or more periods of selection. The rather different situation where estimation must be based upon samples taken from a population before and after a period of selection has been considered in a separate note (Manly, 1975*a*).

A justification for a double exponential fitness function is as follows. Let p_1 be the probability of an individual surviving selection for a unit period of time. Then the probability of survival for time t should be of the form

$$p_t = p_1^{g(t)}$$

where $g(t)$ is a positive, non-decreasing function of t such that $g(0) = 0$ This is because the probability of survival cannot increase as the selection time increases. If conditions remain constant then $g(t)$ will simply be equal to t but in general $g(t)$ will reflect the " rate " at which time is passing from the point of view of survival. Obviously p_1 must always lie in the range $(0, 1)$ and one way to ensure this when p_1 is to be related to the variables $X_1, X_2, ..., X_m$ is to assume that

$$p_1 = \exp\{-\exp(\alpha_0 + \alpha_1 X_1 + ... + \alpha_m X_m)\}.$$

In this case the function $g(t)$ can be incorporated into the constant term α_0 and the fitness function at time t can be written

$$p_t = \exp[-\exp\{\alpha_0(t) + \alpha_1 X_1 + ... + \alpha_m X_m\}] \tag{1}$$

This is what is called the double exponential fitness function. It is a special case of a model proposed by Cox (1972).

Alternative forms that have been suggested for the representation of a

fitness function are the simple quadratic (O'Donald, 1968, 1970, 1971), the " nor-optimal " (Cavalli-Sforza and Bodmer, 1972) and the logistic (Walker and Duncan, 1967). The first two of these suffer from the disadvantage of not necessarily producing fitness estimates in the range (0, 1) while the third has the disadvantage that the relationship between the selection time and the parameters of the function is far from simple.

2. ESTIMATING THE DOUBLE EXPONENTIAL FITNESS FUNCTION

Suppose that the data available for the estimation of a fitness function consists of information on the survival of n individuals in K classes where all the individuals within a class have the same (or approximately the same) values for X_1, X_2, ..., X_m. It is allowable for some or all of the classes to contain only one individual. Suppose also that selection is observed for S intervals of time (t_0, t_1), (t_1, t_2), ..., (t_{S-1}, t_S) and that a_{ij} is the number of survivors in the ith class at the end of the jth interval. Then, providing that $a_{ij-1} \neq 0$, $\hat{p}_{ij} = a_{ij}/a_{ij-1}$ is an estimator of the probability of surviving the jth interval for an individual in the ith class alive at the start of the interval. Estimators of this type are obviously independent if they relate to different classes. Estimators $\hat{p}_{ij}$ and $\hat{p}_{iS}$ are also approximately uncorrelated if $j \neq S$ (Chiang, 1960).

Assuming the double exponential type of fitness function, $\hat{p}_{ij}$ will be estimating

$$p_{ij} = \exp\left\{-\exp\left(\alpha_{0j} + \alpha_1 x_{i1} + \ldots + \alpha_m x_{im}\right)\right\} \qquad (2)$$

where x_{ip} is the value of the variable X_p for individuals in the ith class. By allowing the constants α_{0j} to vary according to the selection interval different interval lengths are allowed. If the number of individuals in each of the K classes is large then equation (2) can be fitted to the data using an empirically weighted multiple regression method (Manly, 1975b, Section 5). However, fitting according to the principle of maximum likelihood is possible using the computer program GLIM that is available from the Numerical Algorithms Group, Banbury Road, Oxford (Nelder and Wedderburn, 1972; Nelder, 1974). GLIM was used for fitting the fitness functions described in the next section of this note.

3. EXAMPLES OF FITNESS FUNCTIONS

(i) The survival of sparrows

Out of 136 exhausted English sparrows taken to the laboratory of Hermon Bumpus after a severe storm in 1898, 72 revived and 64 died. Bumpus realised that this was a situation where the operation of natural selection could possibly be detected and subsequently (Bumpus, 1898) he discussed the relationship between survival and nine characters that he measured on each bird. His conclusion was that there were differences between the survivors and the non-survivors principally related to sex, total length, weight, length of the humerus and length of the femur.

Bumpus published his full set of data with birds classified according to sex and the age of male birds as well as survival. These data have been re-analysed several times since 1898 and, most recently, Johnson et al. (1972)

applied principal component analysis to reach similar conclusions to those of Bumpus whilst O'Donald (1973) examined in some detail the selection related to total length and the length of the humerus.   O'Donald estimated fitness functions based upon the characters of total length, humerus length and a discriminant function (a particular linear combination of total length and humerus length), taken one at a time.   However it seems that no attempt has previously been made to estimate fitness functions of several characters as is done in this note.

Following Johnston *et al.* and O'Donald, weight was excluded from fitness functions on the grounds that a bird's weight was likely to depend to a large extent on its temporary condition.   This leaves the following eight measurements available to " explain " the survival status of Bumpus' birds: total length, alar extent, length of beak and head, length of humerus, length of femur, length of tibio-tarsus, width of skull, and length of keel of sternum. Also a sex variable can be introduced: $X = 0$ for male, 1 for female.   No attempt was made to consider young and old males separately.

For this example there was only one selection period so there is only one α_0 term for the fitness function (2).   No grouping of the individual birds was possible but the computer program GLIM handled the situation quite well except for one problem noted at the foot of table 1 and another problem to do with the measurement of the goodness of fit of fitness functions.   GLIM indicates the goodness of fit of models in terms of what is called the "deviance" from the fitted model, with its corresponding degrees of freedom.   This deviance is actually twice the natural logarithm of the likelihood for the model and is analogous to the residual sum of squares for an ordinary multiple regression (Nelder and Wedderburn, 1972).   For fitting a double exponential function it is normal to regard the deviance as being approximately a chi-squared variate with the stated degrees of freedom and with a significantly high value indicating the lack of fit for the model.   This approximation is however questionable when GLIM is applied to Bumpus' data because when the fitness is assumed to be the same for all 136 birds the deviance is found to be 188·1 with 135 degrees of freedom.   Regarded as a chi-squared variate this is very highly significantly large although this model can hardly fail to fit the data since all that is being done is to estimate a survival probability by a proportion surviving.   The problem seems to be that the chi-squared approximation does not hold for classes consisting only of single individuals.   Nevertheless, a large reduction in the deviance associated with the introduction of certain variables into the fitness function is an indication that fitness was related to these variables.

A summary of the procedure used to fit a fitness function to Bumpus' data is given in table 1.   It will be seen that when all of Bumpus' variables are included in the function then the deviance is 148·6 compared to the value of 188·1 based on a model of constant fitness.   The reduction in the deviance is associated with a loss of nine degrees of freedom and there is therefore an indication that survival was related to at least one of the variables.   This is because if a variable was not related to survival then the introduction of it into the fitness function would be expected to reduce the deviance by approximately one, with a similar reduction in the degrees of freedom.

After fitting various types of fitness function and eliminating any of Bumpus' variables that do not appear to have been related to survival it was finally decided that fitness only appears to have been related to the total

B. F. J. MANLY

Table 1

Double exponential fitness functions fitted to Bumpus' data

Terms included in the fitness function	Deviance	d.f.
Constant term only (all birds with the same fitness)	188·1	135
Sex variable only (different survival for the two sexes)	184·9	134
All of Bumpus' variables plus the sex variable	148·6	126
All of Bumpus' variables and the squares of them plus the sex variable	138·4	118
All of Bumpus' variables with different coefficents allowed for each sex	135·8	118
All of Bumpus' variables and the squares of them with different coefficients for each sex*	?	102
Total length, length of humerus, length of keel of sternum and the sex variable	150·5	131
As for the previous one plus squared terms for the three Bumpus variables included	147·1	128
Total length, length of humerus and the sex variable	153·0	132
Total length and length of humerus with different coefficients for each sex	146·3	130

* For some reason the GLIM program tried to divide by zero when fitting this function and execution was terminated.

length and the length of the humerus of the birds with the relationship being different for the two sexes (*c.f.* final row of table 1). The coefficients for the fitness function were then estimated as follows:

	For males		For females	
Variable	Coef.	Std. err.	Coef.	Std. err.
Constant	− 35·8	10·4	− 6·36	8·61
Total length (mm)	0·402	0·088	0·122	0·073
Humerus length (ins)	− 40·2	10·7	− 18·0	12·1

Because of the nature of equation (2), a negative coefficient indicates that a large value for the variable concerned is associated with a high probability of survival. It therefore appears that short birds with a long humerus had the best chance of surviving the 1898 storm. However the coefficients for the female fitness function are not significantly different from zero when they are compared to their standard errors so that there is some doubt about this conclusion for the females.

Estimated survival probabilities for four of the more extreme birds in Bumpus' sample indicate that the selection caused by the storm was of a very great magnitude:

Total length (mm)	Humerus length (ins)	Sex	Estimated probability of survival
153	0·728	male	0·9711
163	0·689	male	0·0004
153	0·724	female	0·6161
163	0·713	female	0·1354

O'Donald has pointed out in a private communication that the " nor-optimal " fitness functions that he has estimated from Bumpus' data based upon his discriminant function character (O'Donald, 1973, p. 402) give very similar survival probabilities to those in the above table.

(ii) *Over-winter survival of snails*

Table 2 shows the results of an experiment discussed by Cook and O'Donald (1971) involving a large number of *Cepaea nemoralis* that were collected at Cledes, in the Pyrenees, being left without food in an unheated room over the winter of 1968/69. The survivors were counted at the end of six survival periods. Interest centres on the relationship between the shell size, as expressed in the maximum shell diameter, and the survival.

The first step in the analysis of these data involved the fitting of fitness functions as given by equation (2). When only the constants α_{0j} were fitted the deviance was 112·2 with 42 degrees of freedom. The introduction of a term for the snail size reduced this to 68·2 with 41 degrees of freedom while the introduction of a (size)^2 term gave a deviance of 61·5 with 40 degrees of freedom. There was no point in fitting a higher power of size because it was found that even when each of the size classes in table 2 was allowed to have its own survival coefficient the deviance was only reduced to 57·4 with 35 degrees of freedom. The conclusion reached at this point was that the fitness function varied during the full selection period since each of the deviances that have been quoted is significantly large when it is regarded as a chi-squared variate.

The next state in the analysis was to fit a separate fitness function for each of the six survival periods. Taking just linear terms for size this gave a deviance of 49·33 with 36 degrees of freedom. Adding (size)^2 terms then reduced this to 38·73 with 30 degrees of freedom. This latter deviance is not significantly large at the 5 per cent level so that this last model appears to fit the data quite well. It gives the following estimated survival probabilities for the six survival periods:

Max. shell diameter	Selection period					
	1	2	3	4	5	6
21	0·62	0·76	0·61	0·81	0·43	0·44
22	0·82	0·82	0·72	0·79	0·55	0·47
23	0·90	0·85	0·79	0·80	0·64	0·50
24	0·93	0·87	0·83	0·83	0·69	0·51
25	0·95	0·88	0·86	0·87	0·73	0·51
26	0·95	0·89	0·88	0·92	0·74	0·51
27	0·93	0·89	0·89	0·96	0·75	0·49
28	0·90	0·88	0·89	0·98	0·74	0·46

For selection periods 1, 2, 3, 5 and 6 it is estimated that maximum survival occurred for snails with maximum shell diameters in the range 25-28 mm. However, for survival period 4 there is no estimated maximum survival but snails with a maximum shell diameter of about 22 mm are estimated to have the minimum possible survival.

Cook and O'Donald based their analysis of these data on quadratic and " nor-optimal " fitness functions and reached very similar conclusions to those of the present analysis. In particular they noted that the selection during the fourth selection period did not appear to follow the same pattern as the selection in the other periods. One advantage of the analysis based upon the double exponential fitness function is that it was very simple to test the hypothesis that the relationship between fitness and size was constant

for the six selection periods, making allowance for the fact that these intervals were of different lengths.

TABLE 2

The survivors at various times from 1160 Cepaea nemoralis *left over the winter of 1968/69 in an unheated room without food. The snails were collected from Cledes in the Garonne valley a few kilometres south of the border of France and Spain. (Data kindly supplied by Dr L. M. Cook.)*

Maximum shell diameter (mm)	Initial nos. on 29/9/68	Survivors on					
		31/12	7/2	14/3	16/4	21/5	14/6
21	21	15	12	7	4	3	1
22	93	72	57	42	33	14	6
23	255	228	194	152	129	85	44
24	343	326	286	239	192	133	70
25	289	270	239	203	175	132	62
26	128	120	107	99	94	63	36
27	29	29	27	22	21	18	7
28	2	1	1	1	1	1	1

4. REFERENCES

BUMPUS, H. C. 1898. The elimination of the unfit as illustrated by the introduced sparrow *Passer domesticus*. *Biological Lectures*, 209-226. Marine Biol. Lab., Woods Hole.

CAVALLI-SFORZA, L. L., AND BODMER, W. F. 1972. *The Genetics of Human Populations*. Freeman and Co., San Francisco.

CHIANG, C. L. 1960. A stochastic study of the life table and its applications: I. Probability distributions of the biometric functions. *Biometrics, 16,* 618-635.

COOK, L. M., AND O'DONALD, P. 1971. Shell size and natural selection in *Cepaea nemoralis*. In *Ecological Genetics and Evolution*, ed. R. Creed. Blackwell, Oxford.

COX, D. R. 1972. Regression models and life tables. *J. Roy. Statist. Soc. B, 34,* 187-220.

JOHNSON, R. F., NILES, D. M., AND ROHWER, S. A. 1972. Hermon Bumpus and natural selection in the house sparrow *Passer domesticus*. *Evolution, 26,* 20-31.

MANLY, B. F. J. 1975a. The estimation of the fitness function from two samples taken from a population. *Res. Popul. Ecol., 16,* 219-230.

MANLY, B. F. J. 1975b. The measurement of the characteristics of natural selection. *Theor. Pop. Biol., 7,* 288-305.

NELDER, J. A. 1974. *GLIM Manual*. Distributed by the Numerical Algorithms Group, 13 Banbury Rd., Oxford OX2 6NN.

NELDER, J. A., AND WEDDERBURN, R. W. M. 1972. Generalised linear models. *J. Roy. Statist. Soc. A, 135,* 370-384.

O'DONALD, P. 1968. Measuring the intensity of natural selection. *Nature, 220,* 197-198.

O'DONALD, P. 1970. Change of fitness by selection for a quantitative character. *Theor. Pop. Biol., 1,* 219-232.

O'DONALD, P. 1971. Natural selection for quantitative characters. *Heredity, 27,* 137-153.

O'DONALD, P. 1973. A further analysis of Bumpus' data: The intensity of natural selection. *Evolution, 27,* 398-404.

WALKER, S. H., AND DUNCAN, D. B. 1967. Estimation of the probability of an event as a function of several independent variables. *Biometrika, 54,* 167-179.