

**A statistical model for success probabilities  
dependent on group behavior**

Stefania Bartoletti, Daniela S. Monk

Working Paper n. 62, July 2006



**A statistical model for success probabilities  
dependent on group behavior.**

**Stefania Bartoletti**

Ministero del Lavoro e delle Politiche Sociali

Direzione Provinciale del Lavoro di Siena, via della Sapienza, 29 - 53100 Siena

sbartoletti@welfare.gov.it, tel. 0347-8121555, fax: 0577-43448

**Daniela S. Monk**

Department of Biology & Program in Animal Behavior

Indiana University, Bloomington, IN 47405, USA

dmonk@wsu.edu

Dedicated to the memory of Prof. Bernhard Flury (1951 - 1999).

## Abstract

A statistical model is presented to analyze the importance of each of several attributes characterizing a successful individual over other individuals as subjects of a 'choice' situation. The developed model is not a generalized linear model, but it is similar to a polytomous logistic model and treats binary data with covariates. Maximum likelihood estimates are obtained using the Newton Raphson algorithm. We present the application of the model to an analysis of parental care in birds. Specifically we analyze variables related to the successful feeding of one young bird in a nest. We show which conditions the design matrix must satisfy for the information matrix to be positive definite, and under what transformation on the design matrix the model is invariant.

The proposed model is applicable to behavioral studies that analyze the probability of a dichotomous response variable, when several players may affect the outcome.

**keywords:** polytomous logistic regression model, multinomial model, maximum likelihood estimation, score function, information function, Newton Raphson algorithm.

## 1. Introduction

Researchers in behavior frequently are interested in predicting the outcome of social interactions between individuals. In some of these situations, exactly one individual succeeds (wins) and the rest of the individuals do not succeed (lose). Which of the participants turns out to be successful may be a direct consequence of the interaction, as in competition, or may be indirectly determined by an outside party (i. e. a chooser). Examples of these kinds of situations in animal behavior include parental care, where parents choose which offspring to feed and mate selection, where females choose a mate among several rival males (e. g. Andersson and Iwasa 1996, Kilner and Johnstone 1997). Frequently the goal of empirical studies is predicting which individuals are successful, and determining the relative importance of the various attributes of the interactants that contribute to success. However, traditional methods of analyzing these kinds of social interactions generally are not designed to deal with a binary outcome variable, and do not simultaneously account for the effect of multiple interactants and multiple explanatory variables. Here we develop a statistical model to evaluate which characteristics are most important in determining the successful individual among multiple interactants of a group. For these situations, each individual  $j$  participating in the interaction can be described by a vector of covariates  $\mathbf{x}_j$  which consists of attributes that are relevant to the success in the interaction. The outcome of the interaction is dependent not only on the covariate values associated with the successful individual, but also on the covariate values associated with the other individuals involved. The specific example we are analyzing here involves the characteristics of young birds in the nest that determine which one is chosen to be fed by the parent. When a parent bird arrives at its nest, it is greeted by several begging nestlings. At that instant the parent is faced with

the choice of feeding one of many nestlings; this choice depends on the comparison of characteristics of the nestling that ultimately receives food, with the characteristics of all the nestlings that do not receive food. In the application of the model we use a subset of the data obtained from detailed analysis of video tapes from a long-term study on parental care of Mountain Bluebirds (*Sialia currucoides*) (Monk, 1999).

The model developed is not a generalized linear model, but it is similar to a polytomous logistic regression model. The model is based on the comparison of the "intensity" (see section 2) with which the  $j$ -th nestling competes for food, with the sum of the "intensities" of all the nestlings in the same nest.

In section 2 of this paper we describe the model, followed by the derivation of maximum likelihood estimation using the Newton Raphson algorithm in section 3. Section 4 presents the application to the analysis of allocation of parental care among Mountain Bluebirds. In section 5 we show which conditions the design matrix must satisfy for the information matrix to be positive definite, and under what transformation on the design matrix the model is invariant.

## 2. Description of the model

Consider first a single event, that is, one instance of a parent returning to the nest with food for the nestlings. Suppose there are  $k$  nestlings, and only one receives the food. The  $j$ -th nestling has a vector of explanatory variables  $\mathbf{x}_j \in \mathbb{R}^p$ ; these covariates are the coded trait/character values. Let  $\mathbf{y} = (y_1, \dots, y_k)'$  be the response vector of the event, in our case determining which nestling is fed. The vector  $y$  is assumed multinomial, with associated probabilities  $\boldsymbol{\pi} = (\pi_1, \dots, \pi_k)'$ , with  $\sum_{j=1}^k \pi_j = 1$ , and exactly one success. That is, we put  $y_i = 1$  and  $y_j = 0$  for all  $j \neq i$ , if nestling  $i$  receives food.

With a vector  $\boldsymbol{\beta} \in \mathbb{R}^p$  of unknown parameters and a vector

$\mathbf{x}'_j = (x_{j1}, x_{j2}, \dots, x_{jp})' \in \mathbb{R}^p$ , let  $\mathbf{x}'_j \boldsymbol{\beta}$  be a linear predictor, and put

$$\lambda_j = \exp(\mathbf{x}'_j \boldsymbol{\beta}),$$

as in a Poisson regression model with canonical link. The parameter  $\lambda_j$  can be seen as the "sum stimulus" with which the  $j$ -th nestling begs for food. The basic assumption of our model is

$$\pi_j = P(\text{nestling } j \text{ is fed}) = \frac{\lambda_j}{\sum_{h=1}^k \lambda_h}.$$

This is similar to a polytomous logistic regression model. But in contrast to the logistic model we use the same parameter vector  $\boldsymbol{\beta}$  for all categories.

Our model is similar to the Cox's proportional-hazards model (Cox 1972, also in McCullagh and Nelder 1983, chap. 9), and its setup is identical with McFadden's model (McFadden, 1974), of which we were unaware at the time the research for this article was done. The McFadden model is a general procedure for formulating econometric models of population choice behavior from distributions of individual decision rules. However, our terminology and notation seems clearer than McFadden's, and most importantly, the model appears to be so powerful for biological applications that the overlap with McFadden seems to be a minor concern.

Let  $T$  be the total number of events (feedings), and suppose that in the  $t$ -th event there are  $k_t$  nestlings ready to be fed. Let  $\mathbf{y}_t = (y_{1t}, y_{2t}, \dots, y_{k_t t})'$  be the response vector for the  $t$ -th event, and  $\mathbf{x}_{1t}, \mathbf{x}_{2t}, \dots, \mathbf{x}_{k_t t}$  the vectors of covariates (or explanatory variables) associated with the  $k_t$  nestlings, where  $\mathbf{x}_{jt} = (x_{jt1}, x_{jt2}, \dots, x_{jtp})'$ .

Let  $\pi_t = (\pi_{1t}, \pi_{2t}, \dots, \pi_{k_t t})'$  be the vector of probabilities associated with the  $t$ -th event. Then, the multinomial model for the  $t$ -th event is

$$Pr(Y_{1t} = y_{1t}, Y_{2t} = y_{2t}, \dots, Y_{k_t t} = y_{k_t t}) = \prod_{j=1}^{k_t} \pi_{jt}^{y_{jt}},$$

where all  $y_{jt}$  are 0 or 1 and  $\sum_{j=1}^{k_t} y_{jt} = 1$ . Set  $\lambda_{jt} = \exp(\mathbf{x}'_{jt}\boldsymbol{\beta})$ , then our model assumes:

$$\begin{aligned} \pi_{jt} &= P(\text{nestling } j \text{ receives food at time } t) \\ &= \frac{\lambda_{jt}}{C_t} \quad t = 1, \dots, T, j = 1, \dots, k_t, \end{aligned}$$

where  $C_t = \sum_{h=1}^{k_t} \lambda_{ht}$ .

In contrast to the usual definition of regressors of generalized linear models, the vectors  $\mathbf{x}_{jt}$  of our model do not contain a constant, that is there is no intercept term in  $\boldsymbol{\beta}$ . To see why, let

$$\lambda_{jt}^* = \exp(\mathbf{x}'_{jt}\boldsymbol{\beta} + \alpha)$$

for some  $\alpha \in \mathbb{R}$ , then

$$\frac{\lambda_{jt}^*}{\sum_{h=1}^{k_t} \lambda_{ht}^*} = \frac{\exp(\alpha) \exp(\mathbf{x}'_{jt}\boldsymbol{\beta})}{\sum_{h=1}^{k_t} \exp(\alpha) \exp(\mathbf{x}'_{ht}\boldsymbol{\beta})} = \pi_{jt}.$$

That is, a model with intercept would not be identifiable.

### 3. Maximum likelihood estimation

The contribution of the  $t$ -th event to the likelihood function is

$$L_t(\boldsymbol{\beta}/\mathbf{y}_t, \mathbf{X}_t) = \prod_{j=1}^{k_t} \pi_{jt}^{y_{jt}},$$

where  $\mathbf{y}_t$  is the response vector at time  $t$ , and  $\mathbf{X}_t$  is the matrix of the covariates measured on the nestlings at time  $t$ . Assuming independence between events, the likelihood is

$$L(\boldsymbol{\beta}/\mathbf{y}, \mathbf{X}) = \prod_{t=1}^T \prod_{j=1}^{k_t} \pi_{jt}^{y_{jt}}.$$

The logarithm of the contribution to the likelihood of the  $t$ -th event is

$$\begin{aligned} l_t(\boldsymbol{\beta}/\mathbf{y}_t, \mathbf{X}_t) &= \sum_{j=1}^{k_t} y_{jt} \ln \pi_{jt} \\ &= \sum_{j=1}^{k_t} y_{jt} (\mathbf{x}'_{jt} \boldsymbol{\beta} - \ln(C_t)) \\ &= \sum_{j=1}^{k_t} y_{jt} \mathbf{x}'_{jt} \boldsymbol{\beta} - \sum_{j=1}^{k_t} y_{jt} \ln(C_t). \end{aligned}$$

The contribution of the  $t$ -th event to the  $r$ -th score function is

$$\begin{aligned} \frac{\partial l_t(\boldsymbol{\beta})}{\partial \beta_r} &= \sum_{j=1}^{k_t} x_{jtr} y_{jt} - \frac{1}{C_t} \frac{\partial C_t}{\partial \beta_r} \\ &= \sum_{j=1}^{k_t} x_{jtr} y_{jt} - \frac{\sum_{j=1}^{k_t} \exp(\mathbf{x}'_{jt} \boldsymbol{\beta}) x_{jtr}}{C_t} \\ &= \sum_{j=1}^{k_t} x_{jtr} (y_{jt} - \pi_{jt}), \end{aligned}$$

where  $x_{jtr}$  is the  $r$ -th entry of  $\mathbf{x}_{jt}$ . So with respect to the whole parameter vector  $\boldsymbol{\beta}$ , we obtain the contribution of the  $t$ -th event to the score function

$$\mathbf{S}_t(\boldsymbol{\beta}) = \frac{\partial l_t(\boldsymbol{\beta})}{\partial \boldsymbol{\beta}} = \sum_{j=1}^{k_t} \mathbf{x}_{jt} (y_{jt} - \pi_{jt}),$$

and the score function is

$$\mathbf{S}(\boldsymbol{\beta}) = \sum_{t=1}^T \sum_{j=1}^{k_t} \mathbf{x}_{jt} (y_{jt} - \pi_{jt}).$$

The contribution of the  $t$ -th event to the  $(r, q)$ -th information function is

$$-\frac{\partial l_t(\boldsymbol{\beta})}{\partial \beta_r \partial \beta_q} = \sum_{j=1}^{k_t} x_{jtr} \frac{\partial \pi_{jt}}{\partial \beta_q},$$

where

$$\frac{\partial \pi_{jt}}{\partial \beta_q} = \pi_{jt} \left( x_{j tq} - \sum_{j=1}^{k_t} x_{j tq} \pi_{jt} \right),$$

and therefore

$$-\frac{\partial^2 l_t(\boldsymbol{\beta})}{\partial \beta_r \partial \beta_q} = \sum_{j=1}^{k_t} x_{jtr} \pi_{jt} \left( x_{j tq} - \sum_{j=1}^{k_t} x_{j tq} \pi_{jt} \right).$$

Considering the second derivative with respect to the entire  $\boldsymbol{\beta}$  vector, the contribution of the feeding  $t$  to the information function is

$$\mathbf{I}_t(\boldsymbol{\beta}) = -\frac{\partial^2 l_t(\boldsymbol{\beta})}{\partial \boldsymbol{\beta} \partial \boldsymbol{\beta}'} = \sum_{j=1}^{k_t} \mathbf{x}_{jt} \pi_{jt} \left( \mathbf{x}'_{jt} - \sum_{j=1}^{k_t} \mathbf{x}'_{jt} \pi_{jt} \right),$$

and the information function is

$$\mathbf{I}(\boldsymbol{\beta}) = -\frac{\partial^2 l(\boldsymbol{\beta})}{\partial \boldsymbol{\beta} \partial \boldsymbol{\beta}'} = \sum_{t=1}^T \sum_{j=1}^{k_t} \mathbf{x}_{jt} \pi_{jt} \left( \mathbf{x}'_{jt} - \sum_{j=1}^{k_t} \mathbf{x}'_{jt} \pi_{jt} \right).$$

It is straightforward to implement an algorithm for parameter estimation. The algorithm has been implemented in GAUSS (the code may be obtained from the

authors upon request). An explicit maximum likelihood solution is not possible because of the complicated dependence of the score function on the parameters; but numerical maximization of the likelihood function can be done. We use a Newton-Raphson algorithm, starting with an arbitrary initial parameter vector  $\boldsymbol{\beta}^{(0)} \in \mathbb{R}^p$ , and iterating the equation

$$\boldsymbol{\beta}^{(i+1)} = \boldsymbol{\beta}^{(i)} + [\mathbf{I}(\boldsymbol{\beta}^{(i)})]^{-1} \mathbf{S}(\boldsymbol{\beta}^{(i)}) \quad i = 0, 1, 2, \dots$$

until a suitable convergence criterion is met. Our work with this model has shown that the Newton Raphson algorithm usually converges very quickly, typically in at most 6 iterations, starting at  $\boldsymbol{\beta}^{(0)} = \mathbf{0}$ . This contradicts McFadden (1974) who claims that the Newton Raphson algorithm for this model is too slow.

#### 4. Statistical inference

To complete the analysis of the data with our model, we compute also standard errors of the parameters, construct likelihood ratio tests, and perform a residual analysis.

The standard error of each  $\beta_r$ ,  $r = 1, 2, \dots, p$ , is computed taking the square root of the corresponding element on the diagonal of the inverse of the information matrix.

We test for  $H_0 : \boldsymbol{\beta}_{(1)} = \mathbf{0} \in \mathbb{R}^r$  where  $\boldsymbol{\beta} = \begin{pmatrix} \boldsymbol{\beta}^{(1)} \\ \boldsymbol{\beta}^{(2)} \end{pmatrix}$ , with  $\boldsymbol{\beta}^{(1)} = (\beta_1, \beta_2, \dots, \beta_r)'$ ,  $\boldsymbol{\beta}^{(2)} = (\beta_{r+1}, \beta_{r+2}, \dots, \beta_p)'$  and where  $r$  is a number between 1 and  $p$ . Let  $\Omega$  be the unconstrained parameter space, and  $\omega$  the parameter space constrained by the hypothesis  $H_0$ . The log-likelihood ratio statistic for  $H_0$  is

$$\text{LLRS} = 2 \left[ \max_{\boldsymbol{\beta} \in \omega} l(\boldsymbol{\beta}) - \max_{\boldsymbol{\beta} \in \Omega} l(\boldsymbol{\beta}) \right],$$

where  $l(\boldsymbol{\beta}) = \log(L(\boldsymbol{\beta}))$ . If  $H_0$  is true, LLRS is distributed like a chi-square with  $r$  degrees of freedom; we accept the null hypothesis if the  $p$ -value (that is the area that LLRS leaves to the right of the chi-squared distribution with  $r$  degrees of freedom) is greater than a fixed significance level.

Standardized residuals are computed as

$$r_{jt} = \frac{y_{jt} - \hat{\pi}_{jt}}{(\hat{\pi}_{jt}(1 - \hat{\pi}_{jt}))^{1/2}} \quad t = 1, \dots, T, j = 1, \dots, k_t$$

see Cox and Snell (1989), or Dobson (1990). A rough check of the adequacy of the model can be obtained by plotting the residuals and checking for absence of trends.

## 5. Application

For the purpose of illustrating our method, we chose a subset of the data collected on Mountain Bluebirds in Colorado, USA. Results obtained using the entire data set will be reported in Monk (1999). The chosen subset is for parent bluebirds feeding their 13 to 16 day old offspring. We used a total of  $T=360$  events, where each event is a feeding by either the male or female parent. The three covariates we used are also only a subset of all the covariates for which there are data. They are:

$x_{jt1}$ : *relative position*. This variable represents the position of the nestling relative to the feeding parent: rank 1 for nestlings closest to the parent, rank 2 for nestlings where the parent must reach over another nestling in order to feed it, rank 3 for nestlings where parent needs to reach over two other nestlings, etc..

$x_{jt2}$ : *prior feeding*. This indicator variable takes the value 1 when a nestling received food during the preceding event,  $t - 1$ .

$x_{jt3}$ :  $\ln(\text{weight})$ . This variable is the natural logarithm of the weight (in grams) of each nestlings on the day of the event; we use the logarithm of the weight because

the variable we want to consider is the relative weight among the nestlings in the same nest;

For example, 4 events from the considered data file are reported in Table n. 1.

Because there are only 4 nestlings in the example we described, each of the 4 events describes the 4 nestlings with respect to '*relative position*', '*prior feeding*' and '*ln(weight)*'. The first four values of the variable '*prior feeding*' are all 0 because there is no preceding feeding for this nest. For details on the entire data set, please contact the authors.

By running our algorithm for 5 iterations with the chosen data we obtained maximum likelihood estimates for the parameters vector  $\beta$ . We also calculated the standard error for each estimate and the likelihood ratio tests for redundancy of each variable. Our results are summarized in Table n. 2 where LLRS is computed for the redundancy of a single variable each time.

The only non-significant likelihood ratio test is for the variable  $\beta_3$ ; therefore '*ln(weight)*' is considered as redundant. Note that the estimates of  $\beta_1$  and  $\beta_2$  have very low standard errors, indicating that they are fairly stable. The estimate of  $\beta_1$  obtained by our model for this subset of data indicates that Mountain Bluebird parents preferentially feed nestlings that are closer to them. Nestlings where parents have to reach over one or two other chicks receive fewer feedings. The estimate of  $\beta_2$  suggests that when a nestling has not received food during the previous feeding event it is more likely to receive food during the subsequent event. This result suggests that individual feeding events may be dependent on each other, but conditionally on prior feeding the model assumes they are independent; in fact what we need is conditional independence of the response variable given the regressors (see Monk

1999).

To check for the adequacy of the model for this subset of the data we computed standardized residuals and plotted them (Figure I). The absence of trends indicates that the model is adequate.

## 6. Two technical remarks

To confirm the validity of our model it is important to address the two following questions. 1) Under what groups of transformation on the design matrix  $\mathbf{X}$  does the probability remain the same?

Let

$$\mathbf{z}_j = A\mathbf{x}_j + c \iff \mathbf{x}_j = A^{-1}(\mathbf{z}_j - c)$$

be the affine transformation on the design matrix  $\mathbf{X}$ .

The model with  $\mathbf{x}_j$  as a function of  $\mathbf{z}_j$  is

$$\begin{aligned} \pi_{jt} &= \frac{\exp [(\mathbf{z}'_{jt} - c)'(A')^{-1}\beta]}{\sum_{h=1}^{k_t} \exp [(\mathbf{z}'_{ht} - c)'(A')^{-1}\beta]} \\ &= \frac{\exp (\mathbf{z}'_{jt}\gamma + s)}{\sum_{h=1}^{k_t} \exp (\mathbf{z}'_{ht}\gamma + s)}, \end{aligned}$$

where

$$\gamma = (A')^{-1}\beta,$$

$$s = -c'(A')^{-1}\beta.$$

Because of the identifiability condition imposed,  $s$  is simplified and the probability remains the same; only the parameter's vector changes, becoming  $(A')^{-1}\beta$ .

Therefore, the developed model is invariant under nonsingular affine transformation of the design matrix  $\mathbf{X}$ .

2) What conditions on the design matrix  $\mathbf{X}$  lead to the information matrix being positive definite?

For  $\mathbf{I}(\boldsymbol{\beta})$  to be positive definite, for all  $\boldsymbol{\beta} \in \mathbb{R}^p$ , the log-likelihood needs to be concave and therefore have at most one maximum. As  $\mathbf{I}(\boldsymbol{\beta})$  is simply composed of the sum of the  $\mathbf{I}_t(\boldsymbol{\beta})$ , we can consider just one of the contributions to the information matrix and find the conditions on  $\mathbf{X}_t$  so that  $\mathbf{I}_t(\boldsymbol{\beta})$  is positive definite.

After some algebraic manipulations on  $\mathbf{I}_t(\boldsymbol{\beta})$ , we can write:

$$\begin{aligned}\mathbf{I}_t(\boldsymbol{\beta}) &= \sum_{h=1}^{k_t-1} \sum_{j=h+1}^{k_t} \pi_{ht} \pi_{jt} (\mathbf{x}_{ht} - \mathbf{x}_{jt})(\mathbf{x}_{ht} - \mathbf{x}_{jt})' \\ &= \sum_{h=1}^{k_t-1} \sum_{j=h+1}^{k_t} \pi_{ht} \pi_{jt} v_{hjt} v_{hjt}'\end{aligned}$$

where  $v_{hjt} = (\mathbf{x}_{ht} - \mathbf{x}_{jt})$ . In matrix notation the last formula becomes

$$\mathbf{I}_t(\boldsymbol{\beta}) = \mathbf{V}_t' \mathbf{P}_t \mathbf{V}_t,$$

where

$$\mathbf{V}_t = \begin{pmatrix} v'_{12t} \\ v'_{13t} \\ \cdot \\ \cdot \\ \cdot \\ v'_{k_t-1k_t} \end{pmatrix}$$

and

$$\mathbf{P}_t = \text{diag}(\pi_{1t}\pi_{2t}, \dots, \pi_{k_t-1t}\pi_{k_t}).$$

Therefore, the information matrix is positive definite if - the matrix  $\mathbf{P}_t$  has strictly positive diagonal entries (this always holds), - and if the matrix  $\mathbf{V}_t$  has full

column rank, that is if for every covariate there are at least two nestlings with different values of that covariate.

As  $\mathbf{I}(\boldsymbol{\beta})$  is the sum of the  $\mathbf{I}_t(\boldsymbol{\beta})$ , it may have full rank even if some of the  $\mathbf{I}_t(\boldsymbol{\beta})$  are singular.

## Conclusions

In this paper we presented a polytomous regression model that relates multiple covariates to a binary response variable. Our model was designed to analyze the probability of success when several participants are involved in an event and multiple character variables affect the outcome of that event. The model is based on a comparison of the "intensity" (see section 2) with which the  $j$ -th participant competes and the sum of the "intensities" of all the participants. Consequently, our model allows us to determine the probability of success depending on the character variables both of the winning individual and of each of the other participants involved in the same event.

Our model is particularly valuable for the analysis of social behavior because every individual participating in an event likely influences the outcome of that event. Examples include parent-offspring interactions and mate choice, where participants interact in an "arena" and each event results in a single success (Monk 1999). Traditional statistical analysis of social interactions in animal behavior typically are not designed to deal with individual event data, with a binary outcome variable, or to simultaneously account for multiple explanatory variables and multiple participants (Monk and Bartoletti, in prep.).

We developed this model specifically to determine how parent birds make decisions about which offspring to feed. We were interested in predicting the outcome of parent-offspring interactions during feeding events, as well as in determining the

relative importance of offspring attributes that influence parental feeding choices. The model uses character values of all nestlings during each event to calculate success probability. In addition, our approach is designed to evaluate several character traits of nestlings simultaneously and allows us to identify important predictor characters as well as redundant ones. For example, we have shown for Mountain Bluebird parental care that, in the context of this 3-variable analysis, the variable coding for nestling weight had no predictive value, and that the variable describing the position of the nestling in the nest had the most predictive value .

Our model is similar to the Cox and McFadden models (Cox 1972, McFadden 1974), however our notation and terminology is simpler and clearer and thus makes the model readily accessible to biological applications.

Table n. 1

	<i>response</i>	<i>rel. position</i>	<i>prior feeding</i>	<i>weight</i>
<i>nestling1</i>	1	2	0	25.3
<i>nestling2</i>	0	1	0	25.5
<i>nestling3</i>	0	3	0	24.3
<i>nestling4</i>	0	3	0	23.3
<i>nestling1</i>	1	2	1	25.3
<i>nestling2</i>	0	3	0	25.5
<i>nestling3</i>	0	1	0	24.3
<i>nestling4</i>	0	3	0	23.3
<i>nestling1</i>	1	2	0	25.3
<i>nestling2</i>	0	2	0	25.5
<i>nestling3</i>	0	3	1	24.3
<i>nestling4</i>	0	2	0	23.3
<i>nestling1</i>	1	3	0	25.3
<i>nestling2</i>	0	1	0	25.5
<i>nestling3</i>	0	3	0	24.3
<i>nestling4</i>	0	2	1	23.3

**Table n. 2**

<i>parameter</i>	<i>estimate</i>	<i>std. error</i>	LLRS	<i>p-value</i>	<i>degree fr.</i>
$\beta_1$	-2.109	0.145	371.3	0	1
$\beta_2$	-0.360	0.164	5.007	0.025	1
$\beta_3$	0.854	1.358	0.397	0.529	1

## References

- ANDERSSON M, IWASA Y. 1996. Sexual selection. *TREE* 11(2): 53-58.
- BEN-AKIVA M, LERMAN SR. 1993. *Discrete choice analysis*. The MIT Press  
Chambridge, Massachuttes.
- COX DR. 1972. Regression Models and Life Tables (with discussion). *J. R. Statist. Soc., B*, **74**, 187-220.
- COX DR, SNELL EJ. 1989. *Analysis of Binary Data*. Chapman and Hall.
- DOBSON AJ. 1990. *An Introduction to Generalized Linear Models*. Chapman and Hall.
- KILNER R, JOHNSTON RA. 1997. Begging the question: are offspring solicitation behaviours signals of need?. *TREE* 12 (1):11-15.
- MCCULLAGH P, and NELDER JA. 1983. *Generalized Linear Models*. Chapman and Hall.
- MCFADDEN D. 1974. Conditional logit analisys of qualitative choice behavior. In *Frontiers of Econometrics*. P. Zarembka, Academic Press, New York, 105-142.
- MONK DS. 1999, *Differential allocation of parental care in free-ranging Mountain Bluebirds (*Sialia currucoides*)*. PhD Dissertation, Indiana University, Bloomington, IN.