WILD 7250 - Wildlife Population Analysis
www.auburn.edu/~grandjb/wildpop

What are those βs anyway? – Understanding Design Matrix & Odds ratios

References

Hosmer, D.W., and S. Lemeshow. 2000. Applied logistic regression. John Wiley & Sons, Inc. New York.

Review of the design matrix and logit link

Link function used to "link" data to the "real" parameters of interest.

Remember, we are in this class focused on relating our observations back to the demography of the populations we study. Thus, the real parameters of interest are survival, success, recruitment, movement, population growth rate, and distribution (occupancy), as well as the nuisance parameters (e.g., reporting, recapture, and re-sighting rates) included in the maximum likelihood estimators we employ.

The data include not only the response variable (e.g., success or failure), but the additional information we often collect about the individuals or groups that we sample, which may be used to quantify important relationships about those demographic parameters, or to explain the heterogeneity (i.e., variation) among individuals or groups to obtain better estimates of the parameters of interest.

For example, if I am trying to estimate the survival rate of a heavily harvest population of a game species, where there are sex- and age-related differences in natural mortality and harvest, then sex and age are important data in my observations and they can be treated as groups, or covariates in my analysis, but I definitely want to explore the differences in those rates by linking them to the data on sex and age.

The example from the known fate lecture and lab used group, time, and mass as covariates (data). Group and time took on values of 0 or 1, so we refer to them as **indicator** or dummy variables. Covariates such as mass measurements are referred to as **continuous** variables, and measurements or indicators that change over time are referred to as **time varying**.

1. Logit link is the most commonly used link function:

$$S = \frac{e^{(X\beta)}}{1+e^{(X\beta)}} = \frac{\exp(X\beta)}{1+\exp(X\beta)}$$
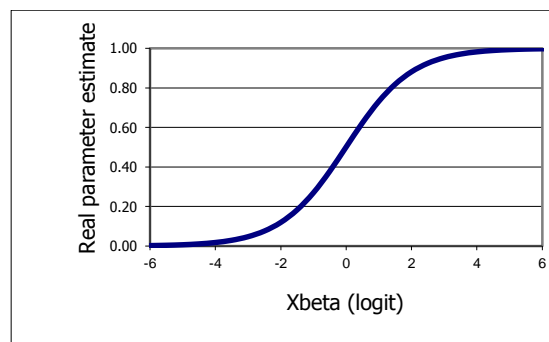
   a. Constrains real parameters to 0-1 interval

WILD 7250 - Wildlife Population Analysis
www.auburn.edu/~grandjb/wildpop

    b. Robust when used in MLE over a relatively wide range of logit values

    c. In the known fate example it is essentially equivalent to using logistic regression.

2. Logit

    a. Equation that describes the relationship of the data and the estimated parameters (βs) in the link function

    b. In MARK this is a linear combination (doesn't have to be)

    This might include 1 parameter

$$x_1\beta_1$$

    or several

$$x_1\beta_1 + x_2\beta_2 \, ,$$

$$x_1\beta_1 + x_2\beta_2 + x_3\beta_3 + x_3^2\beta_4 \, .$$

    c. Often written in generalized matrix notation as $X\beta$

        1) X is the matrix of data (covariates) used to specify the model

            a)   May represent indicator variables (0s and 1s)

            b)   May be values (as in the time variable or group covariate examples)

            c)   May be variable (covariate) names

        2) β are the estimated parameters (i.e., the coefficients in the linear combination)

        3) via matrix multiplication:

$$X\beta = \begin{bmatrix} 1 & x_1 & x_2 & x_3 \\ 1 & x_1 & x_2 & x_3 \\ 1 & x_1 & x_2 & x_3 \\ 1 & x_1 & x_2 & x_3 \end{bmatrix} \begin{bmatrix} \beta \\ \beta_2 \\ \beta_3 \\ \beta_4 \end{bmatrix} = \beta + x_1\beta_2 + x_2\beta_3 + x_3\beta_4$$

3. Design matrix

    a. The most powerful model building tool in MARK

    b. Used to input the logits

        1) Columns are used to specify the βs to be estimated

        2) Each row is used to specify the linear combination of βs that form the logit for one "real" parameter

    c. Example using indicator (dummy) variables – known fate data one group, two time periods.

WILD 7250 - Wildlife Population Analysis
www.auburn.edu/~grandjb/wildpop

Null model (S_i are equal, i.e., estimates a single survival rate):

B1 S Intercept	Parm
1	1:S
1	2:S

$$X = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

$$\beta = [\beta_1]$$

$$X\beta = \begin{bmatrix} 1 \\ 1 \end{bmatrix} [\beta_1] = \begin{bmatrix} \beta_1 \\ \beta_1 \end{bmatrix}$$

So,

$$X_1\beta = \beta_1,$$

and

$$X_2\beta = \beta_1$$

Substituting into the link function:

$$S = \begin{bmatrix} S_1 \\ S_2 \end{bmatrix} = \begin{bmatrix} \dfrac{e^{(\beta_1)}}{1+e^{(\beta_1)}} \\ \dfrac{e^{(\beta_1)}}{1+e^{(\beta_1)}} \end{bmatrix} = \begin{bmatrix} \dfrac{\exp(\beta_1)}{1+\exp(\beta_1)} \\ \dfrac{\exp(\beta_1)}{1+\exp(\beta_1)} \end{bmatrix}.$$

d. Model where the S_i differ between the times (estimates 2 survival rates):

B1 S Intercept	Parm	B2 S t1
1	1:S	1
1	2:S	0

$$X = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix}$$

$$\beta = \begin{bmatrix} \beta_1 \\ \beta_2 \end{bmatrix}$$

$$X\beta = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} \beta_1 \\ \beta_2 \end{bmatrix} = \begin{bmatrix} 1\beta_1 + 1\beta_2 \\ 1\beta_1 + 0\beta_2 \end{bmatrix} = \begin{bmatrix} \beta_1 + \beta_2 \\ \beta_1 \end{bmatrix}$$

So,

WILD 7250 - Wildlife Population Analysis
www.auburn.edu/~grandjb/wildpop

$$X_1\beta = \beta_1 + \beta_2 \, ,$$

and

$$X_2\beta = \beta_1$$

Substituting into the link function:

$$S = \begin{bmatrix} S_1 \\ S_2 \end{bmatrix} = \begin{bmatrix} \dfrac{e^{(\beta_1 + \beta_2)}}{1 + e^{(\beta_1 + \beta_2)}} \\ \dfrac{e^{(\beta_1)}}{1 + e^{(\beta_1)}} \end{bmatrix} = \begin{bmatrix} \dfrac{\exp(\beta_1 + \beta_2)}{1 + \exp(\beta_1 + \beta_2)} \\ \dfrac{\exp(\beta_1)}{1 + \exp(\beta_1)} \end{bmatrix}$$

e. Example using continuous variables (estimates an intercept and trend across intervals).

Expanding the example to include three time periods and estimating the change in survival as a trend over time. Consider the one group example again.

B1 S Intercept	Parm	B2 Time
1	1:S	1
1	2:S	2
1	3:S	3

$$X = \begin{bmatrix} 1 & 1 \\ 1 & 2 \\ 1 & 3 \end{bmatrix}$$

$$\beta = \begin{bmatrix} \beta_1 \\ \beta_2 \end{bmatrix} \quad ,$$

$$X\beta = \begin{bmatrix} 1 & 1 \\ 1 & 2 \\ 1 & 3 \end{bmatrix} \begin{bmatrix} \beta_1 \\ \beta_2 \end{bmatrix} = \begin{bmatrix} 1\beta_1 + 1\beta_2 \\ 1\beta_1 + 2\beta_2 \\ 1\beta_1 + 3\beta_2 \end{bmatrix} = \begin{bmatrix} \beta_1 + \beta_2 \\ \beta_1 + 2\beta_2 \\ \beta_1 + 3\beta_2 \end{bmatrix}$$

So,

$$X_1\beta = \beta_1 + \beta_2 \, ,$$

and

$$X_2\beta = \beta_1 + 2\beta_2 \, ,$$

and

$$X_3\beta = \beta_1 + 3\beta_2$$

Substituting into the link function:

WILD 7250 - Wildlife Population Analysis
www.auburn.edu/~grandjb/wildpop

$$S = \begin{bmatrix} S_1 \\ S_2 \\ S_3 \end{bmatrix} = \begin{bmatrix} \dfrac{e^{(\beta_1+\beta_2)}}{1+e^{(\beta_1+\beta_2)}} \\ \dfrac{e^{(\beta_1+2\beta_2)}}{1+e^{(\beta_1+2\beta_2)}} \\ \dfrac{e^{(\beta_1+3\beta_2)}}{1+e^{(\beta_1+3\beta_2)}} \end{bmatrix} = \begin{bmatrix} \dfrac{\exp(\beta_1+\beta_2)}{1+\exp(\beta_1+\beta_2)} \\ \dfrac{\exp(\beta_1+2\beta_2)}{1+\exp(\beta_1+2\beta_2)} \\ \dfrac{\exp(\beta_1+3\beta_2)}{1+\exp(\beta_1+3\beta_2)} \end{bmatrix}$$

f. Example using variable names (estimates an intercept and trend over covariate values)

Let's introduce a continuous covariate specified in the data something like mass; using the MARK default variable name Var1. This time, it's the name that is constant among individuals, but the values can differ in the input file. Consider the one group example again:

B1 S Intercept	Parm	B2 Var1
1	1:S	Var1
1	2:S	Var1
1	3:S	Var1

$$X = \begin{bmatrix} 1 & \mathrm{Var1} \\ 1 & \mathrm{Var1} \\ 1 & \mathrm{Var1} \end{bmatrix}$$

$$\beta = \begin{bmatrix} \beta_1 \\ \beta_2 \end{bmatrix} \quad ,$$

$$X\beta = \begin{bmatrix} 1 & 1 \\ 1 & 2 \\ 1 & 3 \end{bmatrix} \begin{bmatrix} \beta_1 \\ \beta_2 \end{bmatrix} = \begin{bmatrix} 1\beta_1 + \mathrm{Var1}\beta_2 \\ 1\beta_1 + \mathrm{Var1}\beta_2 \\ 1\beta_1 + \mathrm{Var1}\beta_2 \end{bmatrix} = \begin{bmatrix} \beta_1 + \mathrm{Var1}\beta_2 \\ \beta_1 + \mathrm{Var1}\beta_2 \\ \beta_1 + \mathrm{Var1}\beta_2 \end{bmatrix}$$

So,

$$X_1\beta = \beta_1 + \mathrm{Var1}\beta_2,$$

and

$$X_2\beta = \beta_1 + \mathrm{Var1}\beta_2,$$

and

$$X_3\beta = \beta_1 + \mathrm{Var1}\beta_2$$

Substituting into the link function:

WILD 7250 - Wildlife Population Analysis
www.auburn.edu/~grandjb/wildpop

$$S = \begin{bmatrix} S_1 \\ S_2 \\ S_3 \end{bmatrix} = \begin{bmatrix} \dfrac{e^{(\beta_1 + \text{Var1}\beta_2)}}{1 + e^{(\beta_1 + \text{Var1}\beta_2)}} \\ \dfrac{e^{(\beta_1 + \text{Var1}\beta_2)}}{1 + e^{(\beta_1 + \text{Var1}\beta_2)}} \\ \dfrac{e^{(\beta_1 + \text{Var1}\beta_2)}}{1 + e^{(\beta_1 + \text{Var1}\beta_2)}} \end{bmatrix} = \begin{bmatrix} \dfrac{\exp(\beta_1 + \text{Var1}\beta_2)}{1 + \exp(\beta_1 + \text{Var1}\beta_2)} \\ \dfrac{\exp(\beta_1 + \text{Var1}\beta_2)}{1 + \exp(\beta_1 + \text{Var1}\beta_2)} \\ \dfrac{\exp(\beta_1 + \text{Var1}\beta_2)}{1 + \exp(\beta_1 + \text{Var1}\beta_2)} \end{bmatrix}$$

In this case the logits equations are all the same, and MARK substitutes the appropriate value for the named variable for each observation (capture history) in the input file.

g. Example using additive model

Now building on the examples above, we create an additive model by combining two models for time trend in survival and Var1 (time and Var1). This is called the additive model because these effects are added together in the logit, which is another way of saying that their affect is constant or independent of the other dependent variables (covariates) in the model. So in this example this model estimates the trend in survival over time (β_2) that is consistent over all values of *Var1* and vice-versa.

B1 S Intercept	Parm	B2 Var1	B3 Time
1	1:S	Var1	1
1	2:S	Var1	2
1	3:S	Var1	3

$$X = \begin{bmatrix} 1 & \text{Var1} & 1 \\ 1 & \text{Var1} & 2 \\ 1 & \text{Var1} & 3 \end{bmatrix}$$

$$\beta = \begin{bmatrix} \beta_1 \\ \beta_2 \\ \beta_3 \end{bmatrix} \qquad ,$$

$$X\beta = \begin{bmatrix} 1 & \text{Var1} & 1 \\ 1 & \text{Var1} & 2 \\ 1 & \text{Var1} & 3 \end{bmatrix} \begin{bmatrix} \beta_1 \\ \beta_2 \\ \beta_3 \end{bmatrix} = \begin{bmatrix} 1\beta_1 + \text{Var1}\beta_2 + 1\beta_3 \\ 1\beta_1 + \text{Var1}\beta_2 + 2\beta_3 \\ 1\beta_1 + \text{Var1}\beta_2 + 3\beta_3 \end{bmatrix}$$

So,

$$X_1\beta = 1\beta_1 + \text{Var1}\beta_2 + 1\beta_3 ,$$

and

$$X_2\beta = 1\beta_1 + \text{Var1}\beta_2 + 2\beta_3 ,$$

and

WILD 7250 - Wildlife Population Analysis
www.auburn.edu/~grandjb/wildpop

$$1\beta_1 + \mathrm{Var}1\beta_2 + 3\beta_3$$

Substituting into the link function:

$$S = \begin{bmatrix} S_1 \\ S_2 \\ S_3 \end{bmatrix} = \begin{bmatrix} \dfrac{e^{(1\beta_1 + \mathrm{Var}1\beta_2 + 1\beta_3)}}{1 + e^{(1\beta_1 + \mathrm{Var}1\beta_2 + 1\beta_3)}} \\ \dfrac{e^{(1\beta_1 + \mathrm{Var}1\beta_2 + 1\beta_3)}}{1 + e^{(1\beta_1 + \mathrm{Var}1\beta_2 + 1\beta_3)}} \\ \dfrac{e^{(1\beta_1 + \mathrm{Var}1\beta_2 + 1\beta_3)}}{1 + e^{(1\beta_1 + \mathrm{Var}1\beta_2 + 1\beta_3)}} \end{bmatrix} = \begin{bmatrix} \dfrac{\exp(1\beta_1 + \mathrm{Var}1\beta_2 + 1\beta_3)}{1 + \exp(1\beta_1 + \mathrm{Var}1\beta_2 + 1\beta_3)} \\ \dfrac{\exp(1\beta_1 + \mathrm{Var}1\beta_2 + 1\beta_3)}{1 + \exp(1\beta_1 + \mathrm{Var}1\beta_2 + 1\beta_3)} \\ \dfrac{\exp(1\beta_1 + \mathrm{Var}1\beta_2 + 1\beta_3)}{1 + \exp(1\beta_1 + \mathrm{Var}1\beta_2 + 1\beta_3)} \end{bmatrix}$$

h. Example using interaction

Further building on the examples above, we add an interaction term that estimates the change in the trend in survival over time to change across the values of Var1. We do this by adding another estimated parameter. Within that column in the design matrix we place the product of *Var1* and the value of time. We can do this using the product function in the design matrix (See the Design Matrix Functions in the MARK help file for more on the product function and information on additional functions for use in the design matrix), and the product can be specified in several different ways:

B1 S Intercept	Parm	B2 Var1	B3 Time	B4 Var1*Time
1	1:S	Var1	1	product(col2,col3)
1	2:S	Var1	2	product(col2,col3)
1	3:S	Var1	3	product(col2,col3)

Or

B1 S Intercept	Parm	B2 Var1	B3 Time	B4 Var1*Time
1	1:S	Var1	1	product(Var1,col3)
1	2:S	Var1	2	product(Var1,col3)
1	3:S	Var1	3	product(Var1,col3)

Or

B1 S Intercept	Parm	B2 Var1	B3 Time	B4 Var1*Time
1	1:S	Var1	1	product(Var1,1)
1	2:S	Var1	2	product(Var1,2)
1	3:S	Var1	3	product(Var1,3)

The result is the same:

WILD 7250 - Wildlife Population Analysis
www.auburn.edu/~grandjb/wildpop

$$X = \begin{bmatrix} 1 & \text{Var1} & 1 & \text{Var1} \\ 1 & \text{Var1} & 2 & 2\text{Var1} \\ 1 & \text{Var1} & 3 & 3\text{Var1} \end{bmatrix}$$

$$\beta = = \begin{bmatrix} \beta_1 \\ \beta_2 \\ \beta_3 \\ \beta_4 \end{bmatrix}$$

$$X\beta = \begin{bmatrix} 1 & \text{Var1} & 1 & \text{Var1} \\ 1 & \text{Var1} & 2 & 2\text{Var1} \\ 1 & \text{Var1} & 3 & 3\text{Var1} \end{bmatrix} \begin{bmatrix} \beta_1 \\ \beta_2 \\ \beta_3 \\ \beta_4 \end{bmatrix}'$$

$$= \begin{bmatrix} \beta_1 + \text{Var1}\beta_2 + \beta_3 + \text{Var1}\beta_4 \\ \beta_1 + \text{Var1}\beta_2 + 2\beta_3 + 2\text{Var1}\beta_4 \\ \beta_1 + \text{Var1}\beta_2 + 3\beta_3 + 3\text{Var1}\beta_4 \end{bmatrix}$$

So,

$$X_1\beta = 1\beta_1 + \text{Var1}\beta_2 + 1\beta_3 + 1\text{Var1}\beta_4,$$

and

$$X_2\beta = 1\beta_1 + \text{Var1}\beta_2 + 2\beta_3 + 2\text{Var1}\beta_4,$$

and

$$X_3\beta = 1\beta_1 + \text{Var1}\beta_2 + 3\beta_3 + 3\text{Var1}\beta_4$$

Substituting into the link function:

$$S = \begin{bmatrix} S_1 \\ S_2 \\ S_3 \end{bmatrix} = \begin{bmatrix} \dfrac{e^{(1\beta_1 + \text{Var1}\beta_2 + 1\beta_3 + 1\text{Var1}\beta_4)}}{1 + e^{(1\beta_1 + \text{Var1}\beta_2 + 1\beta_3 + 1\text{Var1}\beta_4)}} \\ \dfrac{e^{(1\beta_1 + \text{Var1}\beta_2 + 2\beta_3 + 2\text{Var1}\beta_4)}}{1 + e^{(1\beta_1 + \text{Var1}\beta_2 + 2\beta_3 + 2\text{Var1}\beta_4)}} \\ \dfrac{e^{(1\beta_1 + \text{Var1}\beta_2 + 3\beta_3 + 3\text{Var1}\beta_4)}}{1 + e^{(1\beta_1 + \text{Var1}\beta_2 + 3\beta_3 + 3\text{Var1}\beta_4)}} \end{bmatrix} = \begin{bmatrix} \dfrac{\exp(1\beta_1 + \text{Var1}\beta_2 + 1\beta_3 + 1\text{Var1}\beta_4)}{1 + \exp(1\beta_1 + \text{Var1}\beta_2 + 1\beta_3 + 1\text{Var1}\beta_4)} \\ \dfrac{\exp(1\beta_1 + \text{Var1}\beta_2 + 2\beta_3 + 2\text{Var1}\beta_4)}{1 + \exp(1\beta_1 + \text{Var1}\beta_2 + 2\beta_3 + 2\text{Var1}\beta_4)} \\ \dfrac{\exp(1\beta_1 + \text{Var1}\beta_2 + 3\beta_3 + 3\text{Var1}\beta_4)}{1 + \exp(1\beta_1 + \text{Var1}\beta_2 + 3\beta_3 + 3\text{Var1}\beta_4)} \end{bmatrix}$$

With any of the above Design Matrices, the deviance, likelihood, number of estimated parameters, estimates of the β_i, and AICc are the same.

i. Interactions among indicator variables

Interactions between indicator polychotomous parameters (more than two possibilities) require each column used in specifying the polychotomous parameter to be multiplied by the columns specifying the other parameter(s).

WILD 7250 - Wildlife Population Analysis
www.auburn.edu/~grandjb/wildpop

For example consider and examples similar to the KM2 example with two groups and only 4 time periods

The default design matrix for two groups (S(g)) is

B1 S Int	Parm	B2 S g1
1	1:S	1
1	2:S	1
1	3:S	1
1	4:S	1
1	5:S	0
1	6:S	0
1	7:S	0
1	8:S	0

The default design matrix for the model for S(t), survival different among times, but not between groups

B1 S Int	B2 S t1	Parm	B3 S t2	B4 S t3
1	1	1:S	0	0
1	0	2:S	1	0
1	0	3:S	0	1
1	0	4:S	0	0
1	1	5:S	0	0
1	0	6:S	1	0
1	0	7:S	0	1
1	0	8:S	0	0

The additive model with group and time effects (S(t+g) (i.e., there is a difference in survival between groups, but it is constant over time).

B1 S Int	B2 S g1	B3 S t1	Parm	B4 S t2	B5 S t3
1	1	1	1:S	0	0
1	1	0	2:S	1	0
1	1	0	3:S	0	1
1	1	0	4:S	0	0
1	0	1	5:S	0	0
1	0	0	6:S	1	0
1	0	0	7:S	0	1
1	0	0	8:S	0	0

Finally, the model with effects of group and time that vary among time periods (S(t*g) is:

WILD 7250 - Wildlife Population Analysis
www.auburn.edu/~grandjb/wildpop

B1 S Int	B2 S g1	B2 S t1	B3 S t2	Parm	B4 S t3	B5 S g1*t1	B6 S g1*t2	B7 S g1*t3
1	1	1	0	1:S	0	1	0	0
1	1	0	1	2:S	0	0	1	0
1	1	0	0	3:S	1	0	0	1
1	1	0	0	4:S	0	0	0	0
1	0	1	0	5:S	0	0	0	0
1	0	0	1	6:S	0	0	0	0
1	0	0	0	7:S	1	0	0	0
1	0	0	0	8:S	0	0	0	0

Odds ratios

The coefficients in the logit (β_i) are interpretable as natural log of the odds ratios. The **odds of success** (survival) are the ratio of the probability of success to the probability of failure, or

$$\frac{S_i}{(1-S_i)}.$$

The **odds ratio** between groups 1 and 2 is then

$$\frac{S_1/(1-S_1)}{S_2/(1-S_2)}.$$

4. Indicator variables

Let's use a simple example of known fate data where 50 individuals were under observation for 2 time periods using telemetry without staggered entry and without censoring. In this example, let 40 individuals survive (fate – 1) the first time period, and 30 survive the second time period.

	Time period	
Fate	1	2
1 (survived)	40	30
0 (died)	10	10
Total	50	40

Survival for period 1 is 40/50 = 0.80, and survival in period 2 is 30/40 = 0.75.

The odds of surviving period 1 are 40/10 = 0.80/0.20 = 4/1 (**ln(4.0) = 1.3863**), and the odds of surviving period 2 are 30/10= 0.75/0.25=3/1 (**ln(3.0) = 1.0986**). The change in odds ratio of survival for the two time period is:

$$\frac{S_1/(1-S_1)}{S_2/(1-S_2)} = \frac{40/10}{30/10} = \frac{4}{3} = 1.33$$

WILD 7250 - Wildlife Population Analysis
www.auburn.edu/~grandjb/wildpop

$$[\beta_1 + \beta_2] - \beta_1 = \beta_2 = 0.287682.$$

Confirming what we illustrated above $e^{\beta_2} = e^{0.287682} = 1.33$. Thus, we could have drawn the same conclusion from examination of the parameters estimated in program MARK. Individuals were 1.33 times as likely to survive period 1 as they were to have survived period 2.

5. Intercept coding

    The reason MARK uses intercept coding in the default design matrix is to allow for the interpretation of the odds ratios among the logits used in estimating the real parameters of interest.

    Consider two design matrices for the known fate Black Duck example S(t):

B1 S Int	B2 S t1	B3 S t2	B4 S t3	Parm	B5 S t4	B6 S t5	B7 S t6	B8 S t7
1	1	0	0	1:S	0	0	0	0
1	0	1	0	2:S	0	0	0	0
1	0	0	1	3:S	0	0	0	0
1	0	0	0	4:S	1	0	0	0
1	0	0	0	5:S	0	1	0	0
1	0	0	0	6:S	0	0	1	0
1	0	0	0	7:S	0	0	0	1
1	0	0	0	8:S	0	0	0	0

and

B1 S t1	B2 S t2	B3 S t3	Parm	B4 S t4	B5 S t5	B6 S t6	B7 S t7	B8 S t8
1	0	0	1:S	0	0	0	0	0
0	1	0	2:S	0	0	0	0	0
0	0	1	3:S	0	0	0	0	0
0	0	0	4:S	1	0	0	0	0
0	0	0	5:S	0	1	0	0	0
0	0	0	6:S	0	0	1	0	0
0	0	0	7:S	0	0	0	1	0
0	0	0	8:S	0	0	0	0	1

The estimates of the real parameters are identical as are the deviance, likelihood, and AICc. However, the estimated parameters in the upper (intercept coded) matrix are interpretable as log odds of effect sizes from the identity matrix below are log odds of success (survival). This concept is easily extended to groups within the data.

6. Continuous dependent covariates

    Given, that we are dealing with linear logit equations in most cases, the β_i associated with continuous dependent variables are interpreted as changes in the log odds ratio per unit change in the value of the variable while holding all other variables constant.

    Using the black duck example again, consider the model S(min<0). In this model, as described in the input file and the Conroy et al. manuscript, min<0 is the number of days

WILD 7250 - Wildlife Population Analysis
www.auburn.edu/~grandjb/wildpop

below freezing during each of the intervals between occasions. These were entered as group covariates in the design matrix. Thus the design matrix and the associated logits are:

B1 S Intercept	Parm	B4 S min<0	Logit
1	1:S	4	$\beta_1 + 4\beta_2$
1	2:S	6	$\beta_1 + 6\beta_2$
1	3:S	7	$\beta_1 + 7\beta_2$
1	4:S	7	$\beta_1 + 7\beta_2$
1	5:S	7	$\beta_1 + 7\beta_2$
1	6:S	6	$\beta_1 + 6\beta_2$
1	7:S	5	$\beta_1 + 5\beta_2$
1	8:S	5	$\beta_1 + 5\beta_2$

The estimates of the β_i were:

Estimated Parameter	Estimate
β_1	6.449301
β_2	-0.60846

Thus we can calculate the logits and the real parameters of interest as:

min<0	Logit	Survival	Odds ratio
3	4.623909	0.990281	
4	4.015445	0.982285	0.544186
5	3.406981	0.967922	0.544186
6	2.798517	0.942596	0.544186
7	2.190052	0.899353	0.544186
8	1.581588	0.829429	0.544186

WILD 7250 - Wildlife Population Analysis
www.auburn.edu/~grandjb/wildpop
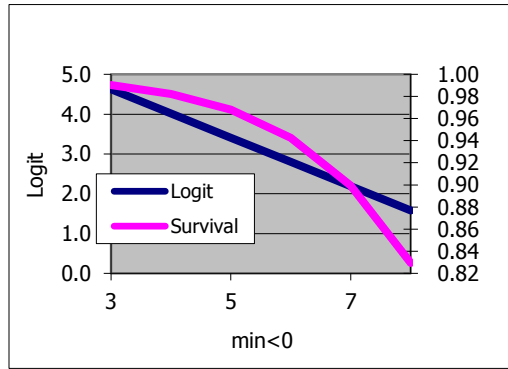
Our interpretation should be that the odds of survival each week are 0.54 ($e^{-0.60846}$) times lower for each additional (one) day below freezing during that week.

7. Multivariable models

   a. Additive models

   When more than one variable is included in the model the interpretation of the odds must include the differences in the distributions of the values for each variable. Consider the Black Duck example again, and the model S(min0+age), where age = 1 for juveniles and age = 0 for adults. The design matrix and the logits are:

B1 S Intercept	Parm	B2 S min<0	B3 S age	Logit
1	1:S	4	age	$\beta_1 + 4\beta_2 + age\beta_3$
1	2:S	6	age	$\beta_1 + 6\beta_2 + age\beta_3$
1	3:S	7	age	$\beta_1 + 7\beta_2 + age\beta_3$
1	4:S	7	age	$\beta_1 + 7\beta_2 + age\beta_3$
1	5:S	7	age	$\beta_1 + 7\beta_2 + age\beta_3$
1	6:S	6	age	$\beta_1 + 6\beta_2 + age\beta_3$
1	7:S	5	age	$\beta_1 + 5\beta_2 + age\beta_3$
1	8:S	5	age	$\beta_1 + 5\beta_2 + age\beta_3$

From the results, the estimated parameters:

Index	Label	β_i	e^{β_i}
1	Intercept	6.524526	681.66
2	min<0	-0.61223	0.54
3	age	-0.16131	0.85

We can calculate the expected survival rates using the link function and the odd ratios between the groups as above

WILD 7250 - Wildlife Population Analysis
www.auburn.edu/~grandjb/wildpop

	Age = 0			Age = 1			ΔOdds ratio between ages
min<0	Logit	Survival	Odds of survival	Logit	Survival	Odds of survival	
3	4.6878	0.9909	108.62	4.5265	0.9893	92.44	0.85
4	4.0756	0.9833	58.89	3.9143	0.9804	50.11	0.85
5	3.4634	0.9696	31.92	3.3021	0.9645	27.17	0.85
6	2.8511	0.9454	17.31	2.6898	0.9364	14.73	0.85
7	2.2389	0.9037	9.38	2.0776	0.8887	7.99	0.85
8	1.6267	0.8357	5.09	1.4654	0.8124	4.33	0.85

Graphing these results we see that the logits are parallel, but the graphs of survival are not despite the fact that the odds ratios remain constant. **Thus, when comparing survival rates for models with more than one variable the values of the parameters at which the relationships are evaluated must be specified explicitly.**
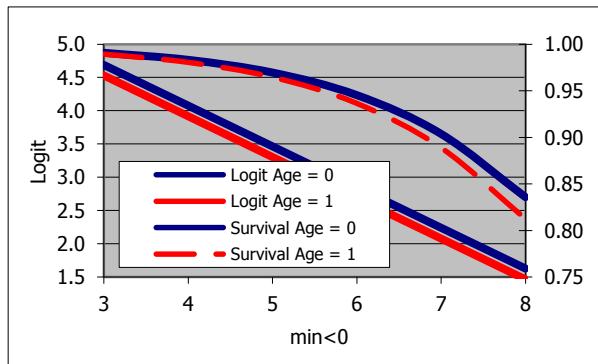
b. Models including interactions

When interactions are estimated in the model even more care must be taken with the interpretation and here's why. Consider the Black Duck example again, and the model S(min0*age), where age = 1 for juveniles and age = 0 for adults. The design matrix and the logits are:

B1 S Intercept	Parm	B2 S min<0	B3 S age	B4 S (min<0)*age		Logit
1	1:S	4	age	product(col2,age)		$\beta_1 + 4\beta_2 + age\beta_3 + 4age\beta_4$
1	2:S	6	age	product(col2,age)		$\beta_1 + 6\beta_2 + age\beta_3 + 6age\beta_4$

WILD 7250 - Wildlife Population Analysis
www.auburn.edu/~grandjb/wildpop

1	3:S	7	age	product(col2,age)	$\beta_1 + 7\beta_2 + age\beta_3 + 7age\beta_4$
1	4:S	7	age	product(col2,age)	$\beta_1 + 7\beta_2 + age\beta_3 + 7age\beta_4$
1	5:S	7	age	product(col2,age)	$\beta_1 + 7\beta_2 + age\beta_3 + 7age\beta_4$
1	6:S	6	age	product(col2,age)	$\beta_1 + 6\beta_2 + age\beta_3 + 6age\beta_4$
1	7:S	5	age	product(col2,age)	$\beta_1 + 5\beta_2 + age\beta_3 + 5age\beta_4$
1	8:S	5	age	product(col2,age)	$\beta_1 + 5\beta_2 + age\beta_3 + 5age\beta_4$

From the results, the estimated parameters:

| Index	Label	Estimate	SE	LCI	UCI
1	Intercept	7.807711	2.676031	2.562691	13.05273
2	min<0	-0.80948	0.405348	-1.60396	-0.015
3	age	-3.14809	3.736018	-10.4707	4.174501
4	min<0*age	0.469423	0.582702	-0.67267	1.61152

We can calculate the expected survival rates using the link function and the odd ratios between the groups as above

	Age = 0			Age = 1			
min<0	Logit	Survival	Odds of survival	Logit	Survival	Odds of survival	Δ Odds between ages
3	5.3793	0.9954	216.86	3.6394	0.9744	38.07	0.18
4	4.5698	0.9897	96.52	3.2994	0.9644	27.10	0.28
5	3.7603	0.9773	42.96	2.9593	0.9507	19.29	0.48
6	2.9508	0.9503	19.12	2.6193	0.9321	13.73	0.72
7	2.1414	0.8949	8.51	2.2792	0.9071	9.77	1.15
8	1.3319	0.7912	3.79	1.9392	0.8743	6.95	1.84

Graphing these results we see that the logits are **NOT** parallel (i.e., they have different slopes). Thus, the odds ratio between age groups is not constant. Thus, **it is even more important, when comparing survival rates for models with more than one variable and interactions, that the values of the parameters at which the relationships are evaluated are specified explicitly.**

WILD 7250 - Wildlife Population Analysis
www.auburn.edu/~grandjb/wildpop
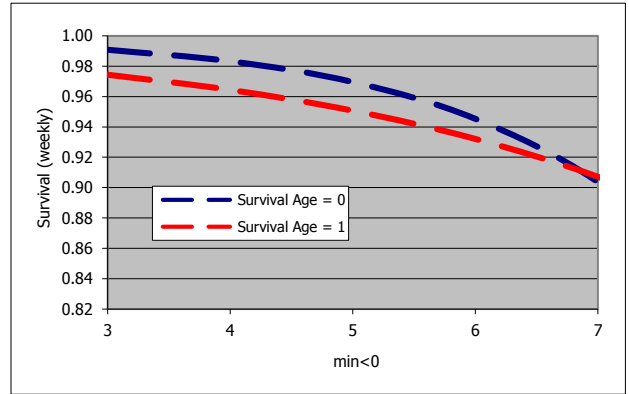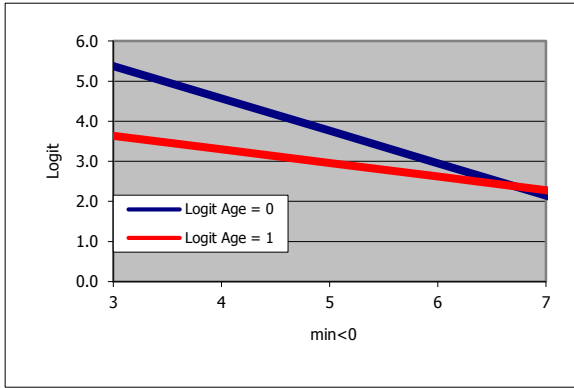
	Age = 0				Age = 1				
			Odds				Odds		
min<0	Logit	Survival	Survival	min<0	Logit	Survival	Survival	Age	min<0
3	5.3793	0.9954	216.86		3.6394	0.9744	38.07	0.18	
4	4.5698	0.9897	96.52	0.45	3.2994	0.9644	27.10	0.28	0.71
5	3.7603	0.9773	42.96	0.45	2.9593	0.9507	19.29	0.45	0.71
6	2.9508	0.9503	19.12	0.45	2.6193	0.9321	13.73	0.72	0.71
7	2.1414	0.8949	8.51	0.45	2.2792	0.9071	9.77	1.15	0.71
8	1.3319	0.7912	3.79	0.45	1.9392	0.8743	6.95	1.84	0.71