**Introduction**

Congress has determined that herds of wild horses and burros are not only a part of the pioneer heritage but have become integral to the ecosystems in which they exist. Because of this, Congress has mandated that these animals be protected and properly managed to protect the numbers and well-being of the herds.

The herds are managed by the Bureau of Land Management, and managed to keep herd numbers at levels that will allow the animals to thrive, but prevent overgrazing. These levels have been termed Appropriate Management Levels (AMLs).

As part of these management efforts, animals are periodically gathered in an area to undergo various procedures such as sterilization, marking, treatments, and others. Excess numbers of animals are auctioned to the public for adoption.

Our intention is to explore the effect of various factors on the successful management of a herd area, and attempt to predict the magnitude of the effects of these chosen predictors. Data collected by the BLM for the fiscal year of 2009 was used for the analysis.

Cases considered were the herd areas managed by the BLM, and success was considered to have occurred when the number of animals was below the AML. Cases where zero animals resided in a managed area were thrown out when considering the years since an animal gathering as a predictor of success, as they would not be representative because they are not a true success because the herds were not preserved in that area, nor would a gather ever be conducted to round up zero animals.

**Methods**

Analysis was conducted using odds ratios, as well as logistic regression, multiple and single. Hypothetical situations with doctored data were also used to more easily demonstrate knowledge of the topics covered in this course.

**The Data**

The data was taken from the Fiscal Year 2009 BLM Wild Horse and Burro Management Data. The proportion of success displayed below was only calculated for the total numbers, and not the data as it is distributed by state. It should be noted that 50 was (somewhat) arbitrarily chosen to represent the number of years since the last gather in areas where a gather had not been recorded, implying that it has been a long time, if ever, since the last gather had been conducted.

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  | NV | | Other Western States | | Total | |  |
| Years Since Last Gather | Number of Areas | AML Success | Number of Areas | AML Success | Number of Areas | AML Success | Proportion |
| 4 | 5 | 2 | 11 | 6 | 16 | 8 | 0.50 |
| 5 | 13 | 6 | 23 | 14 | 36 | 20 | 0.56 |
| 6 | 15 | 5 | 15 | 6 | 30 | 11 | 0.37 |
| 7 | 13 | 3 | 12 | 4 | 25 | 7 | 0.28 |
| 8 | 11 | 2 | 11 | 0 | 22 | 2 | 0.09 |
| 9 | 3 | 2 | 10 | 0 | 13 | 2 | 0.15 |
| 10 | 2 | 1 | 1 | 1 | 3 | 2 | 0.67 |
| 11 | 1 | 1 | 2 | 0 | 3 | 1 | 0.33 |
| 12 | 0 | 0 | 2 | 1 | 2 | 1 | 0.50 |
| 13 | 2 | 1 | 0 | 0 | 2 | 1 | 0.50 |
| 14 | 1 | 0 | 0 | 0 | 1 | 0 | 0.00 |
| 16 | 4 | 2 | 0 | 0 | 4 | 2 | 0.50 |
| 17 | 0 | 0 | 1 | 1 | 1 | 1 | 1.00 |
| 24 | 0 | 0 | 2 | 0 | 2 | 0 | 0.00 |
| 28 | 0 | 0 | 1 | 1 | 1 | 1 | 1.00 |
| 31 | 1 | 1 | 0 | 0 | 1 | 1 | 1.00 |
| 50\* | 8 | 1 | 1 | 1 | 9 | 2 | 0.22 |
| Total | 79 | 27 | 92 | 35 | 171 | 62 |  |

**Odds Ratios**

[ Include some example manipulations ]

*Logistic Regression*

Time Since the Last Gather & The State an Area is Located In

Using logistic regression, it is possible to estimate the likelihood of an event using conditional probabilities. In general, the form of logistic regression is:

Often, when a predictor is binary, it is noted by including a second α parameter, rather than a βX term.

Probabilities can then be predicted for given values using:

The median effective level of a predictor variable in logistic regression, which would indicate the value of the predictor variable required to achieve “the most bang for your buck.” This can be found by the following:

The Models

The First Model: Using only time since the last gather as a predictor



Figure 1: The first logistic model projected onto a plot of the data. While the data is plotted in a manner that differentiates Nevada areas from areas in other states, the model does not account for such a difference.

|  |  |  |  |
| --- | --- | --- | --- |
| Parameter | Value | Std. Error | P-Value |
| α | -.42665 | .29077 | .155 |
| β | -.01457 | .02224 | .518 |
|  |  |  |  |
| AIC | 89.30 | Model P-Value | 0.8904 |

The second model attempted to predict area management success by whether or not the area was located in Nevada or another western state:

Nevada Areas:

Areas in Other States:



Figure 2: The second logistic model projected onto a plot of the data. This model only utilizes two intercept parameters and no slope coefficient. While this model does account for what state the area is located in, it does not account for the time since the last gathering. The silver line represents the effect of the area being located in Nevada; the red line indicates the effects of other states.

|  |  |  |  |
| --- | --- | --- | --- |
| Parameter | Value | Std. Error | P-Value |
| α | -.6554 | .3050 | .042 |
| αOtherState | .1667 | .4115 | .687 |
|  |  |  |  |
| AIC | 89.81 | Model P-Value | 0.9259 |

The third model attempts to predict area management success by both the location of the area and the time in years since the last gathering:

Nevada:

Other States:



Figure 3: The third logistic model projected onto a plot of the data. This model accounts for both the different state levels and the effects of different amounts of time since the last gathering in an area. Two intercepts are used, and two slope coefficients are used. Again, the silver line is for areas in Nevada, red for other states.

|  |  |  |  |
| --- | --- | --- | --- |
| Parameter | Estimate | Std. Error | P-Value |
| α | -.43638 | .43221 | .324 |
| αOtherState | -.09960 | .63501 | .877 |
| β | -.01929 | .02745 | .490 |
| ΒOtherState | .02552 | .05464 | .645 |
|  |  |  |  |
|  |  |  |  |
| AIC | 92.80 | Model P-Value | 0.9988 |

Finally, for knowledge demonstration purposes, we will now consider a hypothetical situation that uses some manipulated data. It should be clear that these results are not valid, as the data is false. The data used was created by removing the 17, 28, and 31 year data points, making it so that “other” states have zero successes for the “no gather” areas and successes in Nevada for the 11 years since the last gather and for “other” states 10 years since the last gather data points were changed to failures. This was done to produce better fits for more easily understood and manipulated models.



Figure 4: The “hypothetical” logistic model projected onto a plot of the manipulated data. This model accounts for both the different state levels and the effects of different amounts of time since the last gathering in an area. Two intercepts are used, and two slope coefficients are used. Again, the silver line is for areas in Nevada, red for other states.

|  |  |  |  |
| --- | --- | --- | --- |
| Parameter | Value | Std. Error | P-Value |
| α | -0.45700 | 0.33688 | 0.175 |
| αOtherState | 3.4035 | 1.06943 | 1.46e-03 |
| β | -0.02736 | 0.02374 | 0.249 |
| ΒOtherState | -0.54483 | 0.16903 | 1.27e-03 |
| AIC | 63.621 | Residual Deviance | 20.135 |

This hypothetical model presents us with an opportunity to evaluate a median effective level that would have some relevance to the data/problem. Using the model as it applies to “other” states, we can see that the median effective level for gathering might be. Remember, the model for this situation is:

So the median effective level would be:

**Conclusions**

Just to be clear, none of these conclusions are valid. The methods applied were done so just for practice and to demonstrate understanding of topics covered in class.