Back-Transforming Estimates

WILD 502 - Jay Rotella

Models for Duck Data: 8 weeks and 1 group

Reconstituting estimates of Survival from the estimated $\beta's$
1. multiply the design matrix by the $\beta's$
2. back-transform the values using the appropriate link function, i.e., the one specified in the analysis

Model S(.)

For this model, the design matrix is 8 rows and a single column of 1's, and a single β is estimated. Thus, survival rate is constrained by the design matrix to be constant across all 8 weeks. Given the data and model structure, $\hat{\beta}_0 = 2.6352908$.

The log-odds of survival in each week is estimated as:

$$\begin{bmatrix}1\\1\\1\\1\\1\\1\\1\\1\end{bmatrix} \cdot \begin{bmatrix}\beta_0\end{bmatrix} = \begin{bmatrix}2.64\\2.64\\2.64\\2.64\\2.64\\2.64\\2.64\\2.64\end{bmatrix}$$

The estimated log-odds value can then be back-transformed using $\frac{exp(\hat{\beta}_0)}{(1+exp(\hat{\beta}_0))}$, which yields 0.9330986 in every week.

Model S(week)

For this model, the design matrix is an 8 x 8 identity matrix, and 8 different $\beta's$ are estimated. Given the data and model structure, the $\hat{\beta}'s$ were estimated as: 3.85, 3.11, 2.97, 1.92, 1.95, 2.12, 3.18, and 3.14.

The log-odds of survival in each week is estimated as follows:

$$\begin{bmatrix}1&0&0&0&0&0&0&0\\0&1&0&0&0&0&0&0\\0&0&1&0&0&0&0&0\\0&0&0&1&0&0&0&0\\0&0&0&0&1&0&0&0\\0&0&0&0&0&1&0&0\\0&0&0&0&0&0&1&0\\0&0&0&0&0&0&0&1\end{bmatrix} \cdot \begin{bmatrix}\beta_0\\\beta_1\\\beta_2\\\beta_3\\\beta_4\\\beta_5\\\beta_6\\\beta_7\end{bmatrix} = \begin{bmatrix}3.85\\3.11\\2.97\\1.92\\1.95\\2.12\\3.18\\3.14\end{bmatrix}$$

Back-transforming the log-odds of survival in each week yields the following weekly survival rates: 0.979, 0.957, 0.951, 0.872, 0.875, 0.893, 0.960, and 0.958, i.e., $\frac{exp(3.85)}{(1+exp(3.85))} = 0.979$, $\frac{exp(3.11)}{(1+exp(3.11))} = 0.957$, $\ldots$, $\frac{exp(3.14)}{(1+exp(3.14))} = 0.958$.

S(time-trend)

For this model, the design matrix is 8 rows and 2 columns (a column of 1's, i.e., the intercept) and a column containing values from 1 to 8. The model imposes a linear time trend on the log-odds of weekly survival. For this model, two $\beta's$ are estimated: $\beta_0 = intercept$ and $\beta_1 = slope$. Thus, the log-odds of survival rate is constrained by the design matrix to follow a linear trend across the 8 weeks. Given the data and model structure, $\hat{\beta}_0 = 3.0652343$ and $\hat{\beta}_1 = -0.1034626$.

The log-odds of survival in each week is estimated as follows:

$$\begin{bmatrix} 1 & 1 \\ 1 & 2 \\ 1 & 3 \\ 1 & 4 \\ 1 & 5 \\ 1 & 6 \\ 1 & 7 \\ 1 & 8 \end{bmatrix} \cdot \begin{bmatrix} \beta_0 \\ \beta_1 \end{bmatrix} = \begin{bmatrix} 3.07 - 0.103 * 1 \\ 3.07 - 0.103 * 2 \\ 3.07 - 0.103 * 3 \\ 3.07 - 0.103 * 4 \\ 3.07 - 0.103 * 5 \\ 3.07 - 0.103 * 6 \\ 3.07 - 0.103 * 7 \\ 3.07 - 0.103 * 8 \end{bmatrix} = \begin{bmatrix} 2.96 \\ 2.86 \\ 2.75 \\ 2.65 \\ 2.55 \\ 2.44 \\ 2.34 \\ 2.24 \end{bmatrix}$$

Notice that the log-odds of survival change by -0.103 (the estimated slope) each week. Back-transforming the log-odds of survival in each week yields the following weekly survival rates: 0.951, 0.946, 0.940, 0.934, 0.927, 0.920, 0.912, and 0.904, i.e., $\frac{exp(2.96)}{(1+exp(2.96))} = 0.951$, $\frac{exp(2.86)}{(1+exp(2.86))} = 0.946$, $\ldots$, $\frac{exp(2.24)}{(1+exp(2.24))} = 0.904$. Those don't change in a perfectly linear pattern because the transformation isn't a linear transformation.

S(temperature)

For this model, the design matrix is 8 rows and 2 columns (a column of 1's) and a column containing a summary of the weekly temperature. For this model, two $\beta's$ are estimated: $\beta_0 = intercept$ and $\beta_1 = slope$. Thus, the log-odds of survival rate is constrained by the design matrix to follow a linear trend with temperature. Given the data and model structure, $\hat{\beta}_0 = 0.7512071$ and $\hat{\beta}_1 = 0.2611014$.

The log-odds of survival in each week is estimated as follows:

$$\begin{bmatrix} 1 & 11.00 \\ 1 & 9.00 \\ 1 & 8.00 \\ 1 & 4.00 \\ 1 & 6.00 \\ 1 & 5.00 \\ 1 & 10.00 \\ 1 & 9.01 \end{bmatrix} \cdot \begin{bmatrix} \beta_0 \\ \beta_1 \end{bmatrix} = \begin{bmatrix} 0.75 + 0.26 * 11.00 \\ 0.75 + 0.26 * 9.00 \\ 0.75 + 0.26 * 8.00 \\ 0.75 + 0.26 * 4.00 \\ 0.75 + 0.26 * 6.00 \\ 0.75 + 0.26 * 5.00 \\ 0.75 + 0.26 * 10.00 \\ 0.75 + 0.26 * 9.01 \end{bmatrix} = \begin{bmatrix} 3.62 \\ 3.10 \\ 2.84 \\ 1.80 \\ 2.32 \\ 2.06 \\ 3.36 \\ 3.10 \end{bmatrix}$$

Notice that the log-odds of survival increase by 0.26 (the estimated slope) for each 1-unit increase in the temperature value (e.g., compare the log-odds of survival in weeks 2 and 3, which had temperature value so 9 and 8, respecitvely: 3.10 - 2.84 = 0.26). Back-transforming the log-odds of survival in each week yields the following weekly survival rates: 0.974, 0.957, 0.945, 0.858, 0.910, 0.887, 0.967, and 0.957, i.e., $\frac{exp(3.62)}{(1+exp(3.62))} = 0.974$, $\frac{exp(3.10)}{(1+exp(3.10))} = 0.957$, $\ldots$, $\frac{exp(3.10)}{(1+exp(3.10))} = 0.957$.

Use R to Make the Calculations

Model S(.)

```r
# make design matrix: a column vector with all ones in it
x = matrix(rep(1, 8), nrow = 8, ncol = 1)

beta = 2.6352908

(Sdot.log_odds = x %*% beta)
```

```
##          [,1]
## [1,] 2.635291
## [2,] 2.635291
## [3,] 2.635291
## [4,] 2.635291
## [5,] 2.635291
## [6,] 2.635291
## [7,] 2.635291
## [8,] 2.635291
```

```r
# 'plogis' command to back-transform to weekly survival rates
(Sdot = plogis(Sdot.log_odds))
```

```
##           [,1]
## [1,] 0.9330986
## [2,] 0.9330986
## [3,] 0.9330986
## [4,] 0.9330986
## [5,] 0.9330986
## [6,] 0.9330986
## [7,] 0.9330986
## [8,] 0.9330986
```

Model S(week)

```r
# make design matrix (an identity matrix) with 'diag' function
x=diag(8)

# declare beta-hat values
beta=c(3.8501476,3.1135152,2.9704144,1.9169226,
1.9459101,2.1202635,3.1780540,3.1354941)

(St.log_odds=x%*%beta)
```

```
##          [,1]
## [1,] 3.850148
## [2,] 3.113515
## [3,] 2.970414
## [4,] 1.916923
## [5,] 1.945910
## [6,] 2.120264
## [7,] 3.178054
## [8,] 3.135494
```

```r
(St=plogis(St.log_odds))
```

```
##           [,1]
## [1,] 0.9791667
## [2,] 0.9574468
## [3,] 0.9512195
## [4,] 0.8717949
## [5,] 0.8750000
## [6,] 0.8928571
## [7,] 0.9600000
## [8,] 0.9583333
```

S(time-trend)

```r
# construct the 2-column design matrix
(x = matrix(c(rep(1, 8), 1:8), nrow = 8, ncol = 2, byrow = FALSE))
```

```
##      [,1] [,2]
## [1,]    1    1
## [2,]    1    2
## [3,]    1    3
## [4,]    1    4
## [5,]    1    5
## [6,]    1    6
## [7,]    1    7
## [8,]    1    8
```

```r
# declare beta-hat values
beta = c(3.0652344, -0.1034626)

(ST.log_odds = x %*% beta)
```

```
##          [,1]
## [1,] 2.961772
## [2,] 2.858309
## [3,] 2.754847
## [4,] 2.651384
## [5,] 2.547921
## [6,] 2.444459
## [7,] 2.340996
## [8,] 2.237534
```

```r
(ST = plogis(ST.log_odds))
```

```
##           [,1]
## [1,] 0.9508169
## [2,] 0.9457466
## [3,] 0.9401865
## [4,] 0.9340962
## [5,] 0.9274337
## [6,] 0.9201553
## [7,] 0.9122159
## [8,] 0.9035698
```

S(temperature)

```r
# make design matrix: 2 columns (ones in 1st, temps in 2nd)
temps = c(11, 9, 8, 4, 6, 5, 10, 9.01)
(x = matrix(c(rep(1, 8), temps), nrow = 8, ncol = 2, byrow = FALSE))
```

```
##      [,1]  [,2]
## [1,]    1 11.00
## [2,]    1  9.00
## [3,]    1  8.00
## [4,]    1  4.00
## [5,]    1  6.00
## [6,]    1  5.00
## [7,]    1 10.00
## [8,]    1  9.01
```

```r
# declare beta-hat values
beta = c(0.7509425, 0.2611643)
(Stemp.log_odds = x %*% beta)
```

```
##          [,1]
## [1,] 3.623750
## [2,] 3.101421
## [3,] 2.840257
## [4,] 1.795600
## [5,] 2.317928
## [6,] 2.056764
## [7,] 3.362585
## [8,] 3.104033
```

```r
(Stemp = plogis(Stemp.log_odds))
```

```
##           [,1]
## [1,] 0.9740110
## [2,] 0.9569513
## [3,] 0.9448129
## [4,] 0.8576124
## [5,] 0.9103510
## [6,] 0.8866293
## [7,] 0.9665146
## [8,] 0.9570588
```

What about Measures of Uncertainty?

We still need to consider how one propagates the uncertainty in the $\hat{\beta}'s$ into estimates of survival rate. There are several ways of doing so. As you've seen, Program MARK has some useful tools for putting uncertainty on estimates of the real parameters of interest (e.g., survival rates) given the uncertainty in the $\hat{\beta}'s$ as measured by the variance-covariance matrix for the $\hat{\beta}'s$. We'll discuss how Program MARK does this and show you how you can implement the delta method, which is one commonly used approach, in R in the coming weeks.