

Logistic Regression in MARK

Why MARK for this? – unified tool for the semester

Why do input files look different for labs 1 & 2?

Lab 1 – only interested in time variation, not in individual variation

```
known fate group=1;
```

```
48 1;
```

```
47 2;
```

```
41 2;
```

```
39 5;
```

```
32 4;
```

```
28 3;
```

```
25 1;
```

```
24 1;
```

```
/* Fawn survival for 1 interval,
```

```
Comments provide (1) radio frequency and (2) days lived.
```

```
Live-dead encounter history for 1 occasion comes next.
```

```
Number of animals with this encounter history comes after encounter history.
```

```
Individual Covariates are:
```

```
1) Area: 0=Control 1=Treatment
```

```
2) Sex: 0=Female 1=Male
```

```
3) Mass (kg)
```

```
4) Length (cm) */
```

```
/* 191.62 10 */ 11 1 0 34 126 ;
```

```
/* 192.6 11 */ 11 1 0 33.4 128.5 ;
```

```
/* 191.64 12 */ 11 1 0 29.9 130 ;
```

```
/* 192.79 16 */ 11 1 0 31 113 ;
```

```
/* 192.22 16 */ 11 1 0 26.7 116 ;
```

```
/* 192.53 13 */ 11 1 0 23.9 124 ;
```

```
...
```

```
/* 191.075 . */ 10 1 1 34.2 125.5 ;
```

```
/* 191.61 . */ 10 1 0 31.8 122 ;
```

```
/* 191.63 . */ 10 1 0 37 125.5 ;
```

```
/* 192.15 . */ 10 1 0 31 118 ;
```

```
/* 192.16 . */ 10 1 0 34.5 121 ;
```

```
/* 192.17 . */ 10 1 0 33 127.5 ;
```

```
/* 192.18 . */ 10 1 0 35.5 128 ;
```

```
/* 192.19 . */ 10 1 0 32.5 124 ;
```

```
/* 192.32 . */ 10 1 1 33.9 125 ;
```

```
/* 192.34 . */ 10 1 1 37.5 128 ;
```

```
...
```

```
/* 192.79 119 */ 11 1 0 1 40 133 ;
```

```
/* 192.09 119 */ 11 1 0 0 28 118.5 ;
```

```
/* 192.71 130 */ 11 1 0 1 38.6 131 ;
```

What if want time and individual variation?

Can do this quite easily by combining the 2 approaches, e.g., with black ducks.

```
/* Conroy black duck radiotracking data,
   Encounter occasions=8, groups=1, individual covariates=4,
   individual covariate names = Age (0=subadult, 1=adult),
   Weight (kg), Wing Length (cm), and Condition Index. */

/* 01 */ 1100000000000000 1 1 1.16 27.7 4.19;
/* 04 */ 1011000000000000 1 0 1.16 26.4 4.39;
/* 05 */ 1011000000000000 1 1 1.08 26.7 4.04;
/* 06 */ 1010000000000000 1 0 1.12 26.2 4.27;
/* 07 */ 1010000000000000 1 1 1.14 27.7 4.11;
/* 08 */ 1010110000000000 1 1 1.20 28.3 4.24;
/* 09 */ 1010000000000000 1 1 1.10 26.4 4.17;
...

/* 29 */ 1010101010101010 1 0 1.18 25.9 4.56;
/* 30 */ 1010101010101010 1 0 1.07 26.7 4.01;
/* 31 */ 1010101010101010 1 1 1.26 26.9 4.68;
/* 32 */ 1010101010101010 1 0 1.27 27.6 4.60;
/* 33 */ 1010101010101010 1 0 1.08 26.0 4.15;
/* 34 */ 1010101010101010 1 1 1.11 27.0 4.11;
/* 35 */ 1010101010101010 1 0 1.15 27.1 4.24;
...

/* 43 */ 1010101010101010 1 0 1.15 26.3 4.37;
/* 44 */ 1010101010101010 1 1 1.27 27.0 4.70;
/* 45 */ 1010101010101010 1 1 1.37 27.5 4.98;
/* 46 */ 1010101010101010 1 1 1.22 26.5 4.60;
/* 47 */ 1010101010101010 1 0 1.22 26.8 4.55;
/* 48 */ 1010101010101010 1 0 1.14 26.2 4.35;
/* 49 */ 1010101010101010 1 0 1.14 27.0 4.22;
/* 50 */ 1010101010101010 1 0 1.12 27.4 4.09;
```

To see how to a set of models run this way, start MARK and open "BLCKDUCK.DBF" - it can be found in "C:\Program Files\MARK\Examples" in a standard installation.

When we did Lab 1, we only focused on time variation in survival rates. Thus, it was reasonable to be interested in model-averaged weekly survival estimates.

Once we add individual covariates, things get more complicated: one has to think about weekly survival rate for a bird with what characteristics. For now, we won't get into model averaging when individual covariates are involved - we will deal more with this later in the semester though.

Another issue comes up when working with a design matrix with more than 1 column: how do I put a variance on the log-odds of survival when there are multiple beta's in the model? It turns out that there are several ways of doing this and we'll learn about them later in the course. For now, realize that the trick is dealing with the fact that we have a sampling variance associated with each of the beta's in the model and a sampling covariance associated with each pair of beta's in the model. So, to put a variance on the log-odds of survival, we need to incorporate those uncertainties. MARK does this using the delta method, and we'll learn how that method works later in the course. It is

also used in lots of other regression modeling when you have multiple beta's in a model. If you want to implement the delta method before we discuss how to it works, here's some R code for an example from Lab 2. Note that we first estimate SE's & CI's on the log-odds scale and then back-transform those CI limits to the real parameter scale.

```
# example of delta method for length model - fawns data
library(emdbook) # for delta method function

b0=-10.635836 # intercept
b1=0.0831520 # beta for length

sigma=matrix(c(
22.8640286467, -0.1845438929,
-0.1845438929,  0.0014919721),2,2)

# Provide length values over range of data
Length=seq(from=108, to=135.5, by=0.5)

# Calculate ln odds of Survival rate for any length
ln_odds_S=b0+b1*Length

# Estimate the Standard Errors for ln odds of S
se_ln_odds_S=sqrt(deltavar(b0+b1*Length,meanval=c(b0=-10.635836,b1=0.0831520),Sigma=sigma))

# Estimate the 95% confidence limits for ln odds of S
lcl_ln_odds_S=ln_odds_S - 1.96*se_ln_odds_S
ucl_ln_odds_S=ln_odds_S + 1.96*se_ln_odds_S

# back transform to real parameters
S=plogis(ln_odds_S)
lcl_S=plogis(lcl_ln_odds_S)
ucl_S=plogis(ucl_ln_odds_S)

# Estimate the Standard Errors for S
se_S=sqrt(deltavar(exp(b0+b1*Length)/(1 + exp(b0+b1*Length)),
  meanval=c(b0=b0,b1=b1),Sigma=sigma))

plot(Length,S,'l',ylim=c(0,1))
lines(Length,lcl_S,'l',lty=3)
lines(Length,ucl_S,'l',lty=3)
```

