

# Update 3 March 2016

*Erin Feichtinger*

*March 3, 2016*

## Update of Research and Studies March 3 2016

### Survival Analyses of Florida Scrub-Jays

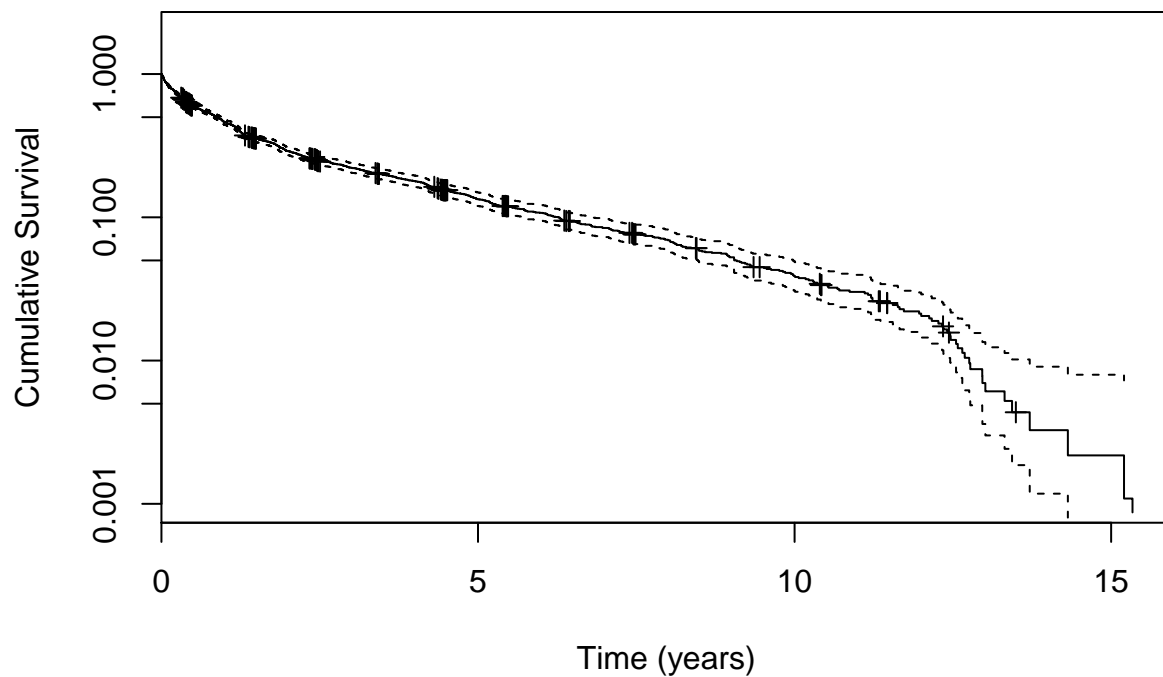
I used two subsets of the jay data to fit Kaplan-Meier survival curves, Cox Proportional Hazard models and Accelerated Life or Failure Time models. The two subsets of data are 1) all-known age birds and 2) breeders of both known and unknown age from 1981 to 2015. I fit KM curves for both sexes combined and males and females separately for both subsets of data. Following are the Kaplan Meier Curves.

```
## Loading required package: Matrix
```

```
## Loading required package: quadprog
```

```
## Warning in xy.coords(x, y, xlabel, ylabel, log): 1 y value <= 0 omitted  
## from logarithmic plot
```

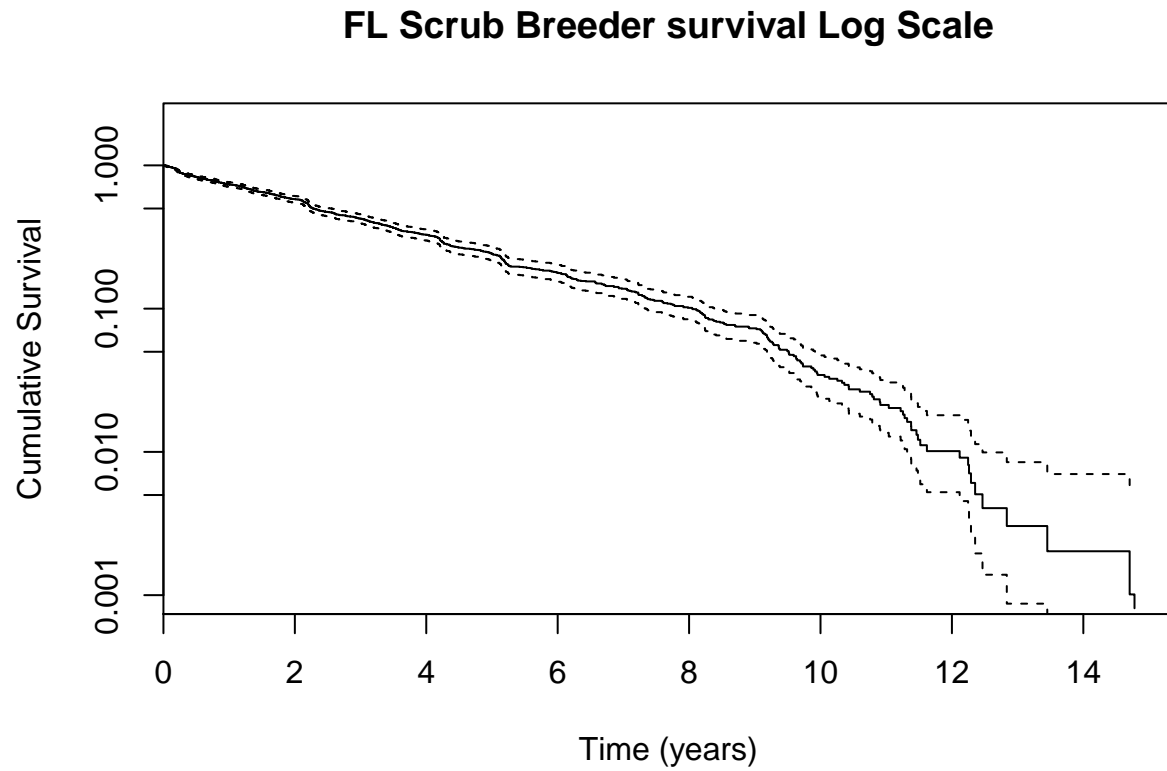
### Survival of Known Age Birds – Log Scale



```
## [1] 493
```

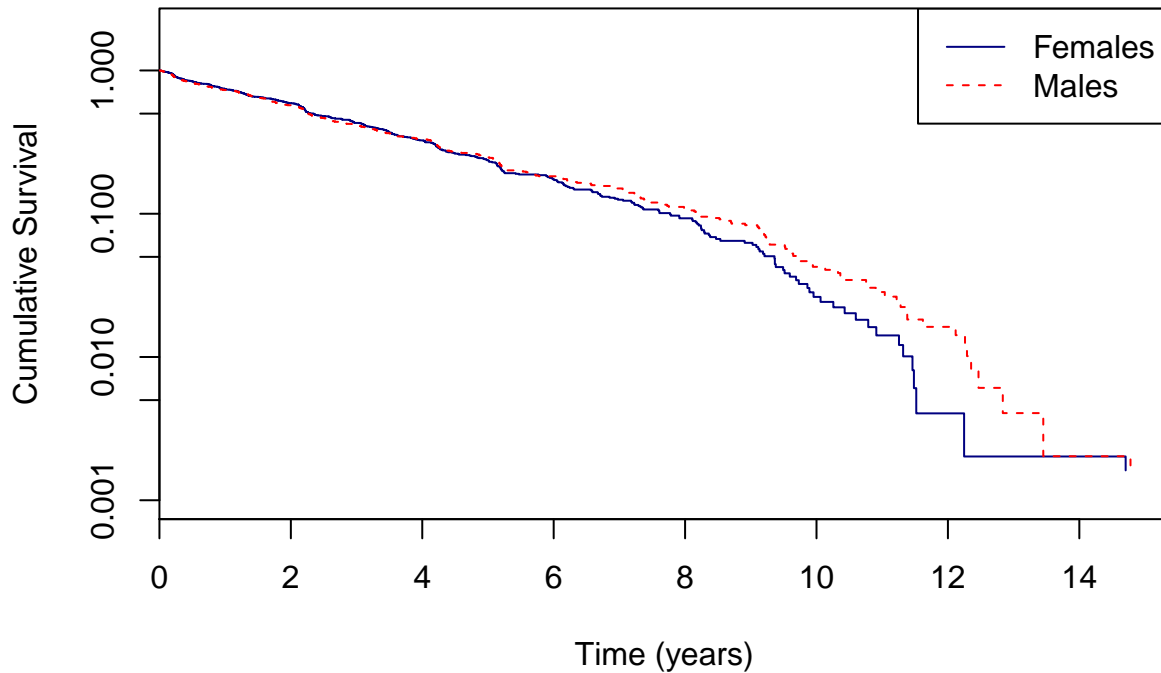
```
## [1] 495
```

```
## Warning in xy.coords(x, y, xlabel, ylabel, log): 1 y value <= 0 omitted  
## from logarithmic plot
```



```
## Warning in xy.coords(x, y, xlabel, ylabel, log): 1 y value <= 0 omitted  
## from logarithmic plot
```

## Survival by Sex Log Scale



### Semi-parametric and parametric models for breeders

I have fit some basic Cox and AFT models to both data sets with combinations of three predictors, year of first breeding, age at first breeding, and sex. I found that sex is not a significant predictor for any Cox model that includes it alone or with other covariates for breeders. Age at first breeding is significant, but  $R^2$  is really low. I coded year of first breeding as a factor so there are 34 levels (34 years). Some coefficients for year, including 2011-2015, are significant and  $R^2 = 0.23$ . A Cox model with both year and age at first breeding has an associated  $R^2 = 0.24$ , nearly identical to the model with year alone. Although the coefficient estimate for age at first breeding is significant that results in an increase in hazard, it does not seem to add more information to the model that includes year.

For the AFT models, sex is significant but I don't think it is meaningful although I'm not sure how to interpret the model output. The Cox output is a little easier for me to interpret at this time. The sign of the coefficient is opposite to the estimate for the Cox model for sex. Age at first breeding is significant but again, I don't know how much of survival times it explains. Year is also significant with a negative coefficient (it was positive in the Cox model). I used `extractAIC()` to get the AIC scores for a set of models that included sex, age at first breeding in various combinations. The model with the lowest AIC for this set was the one with all three covariates: year, age at first breeding and sex. Likewise, for the Cox models, the model with all 3 covariates had a very similar AIC score to the model with year and age at first breeding. So far, it seems that year is important while the other two covariates not as much.

## **Next Steps of Survival Analyses**

I need to get group size to add to the subsets of data. However there are a few issues here. First, I have been trying to figure out how to add group size to the data set in the jay database but it is not straightforward. I haven't figured out how to connect two existing queries that have all of the necessary information.

Also, I have some questions about how to use group size in the models. For the time-varying Cox models, every individual can have multiple rows so I can include group size at the April census each year. However, I'm not sure how to go about this for the simpler Cox models and the AFT models which require the data to be set up in a way so each individual has just one record. How can I use group size as a covariate here? The only options that I can think of are at the year of first breeding (which would likely be zero) or year a bird fledged, or use group size the year of death. I don't think either of those options are a useful or informative option.

The other covariates of interest are time since fire in the territory, acorn abundance and pedigree. Acorn abundance is already in a useable format, short of centering. I have a yearly summary on each territory of acorns. The fire data and the pedigree are a little more challenging. We have talked about fire in the past and I need to remind myself how we planned to do this. Contacting Reed and Becky is a good idea. As far as the pedigree, I attempted to code it this summer but I ran into some syntax issues with correctly identifying the founders of the pedigree so it wil work. I haven't yet solved the issue but I looked at the script file and the documentation on the kinship2 package this week to remind myself how to go about solving this problem.

## **Growth Rate Heterogeneity**

Bruce read the manuscript and sent back the draft with his comments and feedback. As of today (Thursday), I haven't had a chance to carefully read through it. The plan is to do that over the next few days and be prepared to talk about it with Gordon. Additionally, we have discussed having a skype meeting with Bruce so all three of us can talk about it. I imagine that to be the next step. After we have done that, I think fixing the figures and making them really nice is a good idea.

## **Other happenings and plans for the next week and through spring break**

I had several assignments to grade this past week so I could submit my midterm grades. I have 92 more assignments to grade over spring break. I am taking an R course this spring and I have to devote some time to that over the next week and during spring break. I have a proposal for a package and I am working on writing up a more detailed report. Once I have done that, I will discuss my idea with Gordon.

1. Work on the next steps for the survival analyses
2. Read through manuscript and think about the feedback and suggestions from both Bruce and Gordon
3. Work on my R package
4. Visit Archbold for a day or two during spring break