

Survival

Erin Feichtinger

January 25, 2016

I have done some “preliminary” work if you will, last summer and fall. I have extracted data on the jays from the database with the information needed for basic survival models. From there, I set up a file with birds that became breeders. The start of the interval for this stage is the year a bird was first recorded as a breeder and the end is April 2015 (at least it was then, I can always add more months from 2015). I chose April because that is when individuals move from juvenile/independent young to one year old (adult helper). A bird can become a breeder earlier in spring and nestlings fledge all spring. Back to breeders for a moment, in April 2015, they are either dead or still alive (right censored).

The starting point now is to fit survival models to the data. I made some survival objects using the Survival Package in R with the Surv function. I think a good place to start is to have all of the birds in one data set not split up by stage. Does doing this step make it justifiable to split the population up into stages for modeling survival? My intuition is to start here and see if we can adequately describe survival with simple models. I predict that we will have justification to split the birds up by stage because the survival probability differs by stage.

First, I’ll use the Kaplan Meier estimate of survival. I’ve done this before but I’m not really sure what it tells me that’s useful here (besides showing the cumulative survival of the birds in the sample). I followed the example laid out in Gordon’s Intro to Event Times from September 2009 on the cactus finch first. The plot of the curve follows. I used the summary() function and the output was crazy (lots of individuals). The first thing that I notice is that there are a lot of events (i.e. deaths!!) in the first part of the time period which runs from Jan 1981 to October 2015 in this case. The survival starts out high but drops quickly. This is consistent with what we know about the jay biology. The young ones have a higher mortality rate than adults. But what can this tell us about the need to split up the birds by stage? The number of deaths in each interval declines over time.

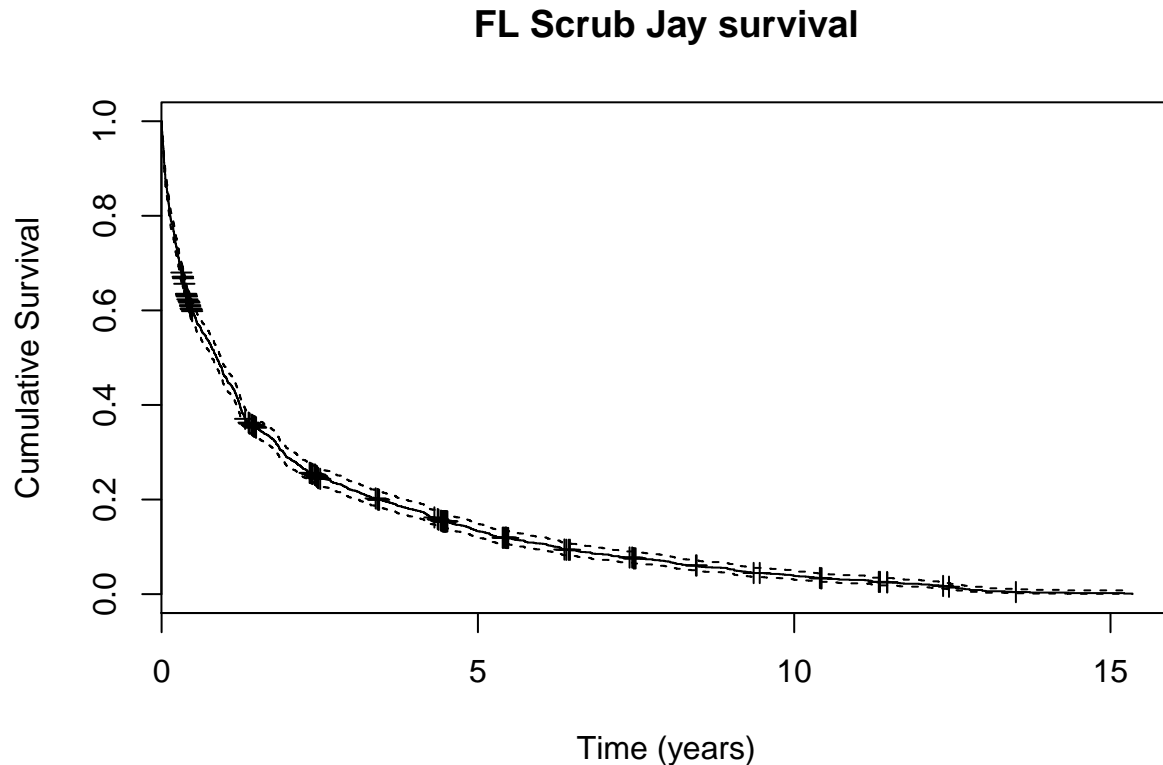
```
require(survival)
```

```
## Loading required package: survival
```

```
setwd("C:/Users/efeichtinger/Reports")
birds <- read.csv("Erin_Surv_All.csv")
birds2 <- birds[!duplicated(birds),]
colnames(birds2)[6] <- "LastObsDate"
birds2$FldgDate <- as.Date(birds2$FldgDate, format = "%m/%d/%Y")
birds2$LastObsDate <- as.Date(birds2$LastObsDate, format = "%m/%d/%Y")
date.diff <- birds2$LastObsDate - birds2$FldgDate
birds2["days"] <- date.diff
birds2$days <- as.numeric(birds2$days)
birds2["yrs"] <- birds2$days/365
birds2["censorship"] <- 1
birds2$censorship[which(birds2$LastObsDate=="2015-10-14")] <- 0
birds2$FldgDate <- as.numeric(birds2$FldgDate)
birds2$LastObsDate <- as.numeric(birds2$LastObsDate)
birds2$days <- as.numeric(birds2$days)
birds2$yrs <- as.numeric(birds2$yrs)

survobj <- Surv(birds2$yrs, birds2$censorship)
```

```
jay.lifetab <- survfit(survobj~1)
jay.fit <- plot(jay.lifetab, xlab = "Time (years)",
  ylab = "Cumulative Survival", main = "FL Scrub Jay survival")
```



Why does this plot not look like the one in Gordon's report? It doesn't have the stepwise pattern. Or does it and I can't tell because there are a greater number of samples?

Now I'm not really sure where to go next. Gordon's advice was to start fitting basic models. I still don't know what this means even though we had talked about it. The exponential model? I don't have the covariates in the data set currently (except for sex and yr born). They are accessible but I need to add them in. I started this way to get a feel for the survival modeling process with a few as covariates as possible.

According to P. Keli (2015), when the hazard at time t is independent of age, the prob density for the proportion of the cohort at time t declines exponentially... a constant hazard.

I tried fitting an exponential model with no covariates but I got an error: Error in survreg(survobj ~ 1, dist = "weibull") : Invalid survival times for this distribution In addition: Warning messages: 1: In log(dlist\$dtrans(Y[exactsurv, 1])) : NaNs produced 2: In log(y) : NaNs produced

Same for the Weibull model

References

Fox, G. A. 2001. Failure-time analysis in Design and Analysis of Ecological Experiments.

Fox, G. A. 2009. Introduction to failure time data & censored regression.

Keli, P. 2015. Survival analysis: basic terms, the exponential model, censoring, examples in R and JAGS.

Rodriguez, G. 2007. Survival Models.

Slud, E. V. Survival Analysis Models & Statistical Methods - Short Course.