**Preliminary Exam Questions: Dr. Dinsmore**

Marty Simonson

March 27, 2020

4. Please complete a short analysis of the seasonal survival of a fish (species is irrelevant) in a hypothetical lake system. I have attached two files (.dbf and .ftp) that are formatted for Program MARK. I am interested in seeing your approach to this type of analysis, with less emphasis on a “right” answer. Also realize that this is a “real” analysis in the sense that the data might not be perfect, all parameters might not be estimable, etc. Please begin by providing a *brief* justification for the factors that you considered in your analyses, and your general approach to building models for consideration. Then run an appropriate set of models to answer the questions below. I’ll want to see both of your MARK files, so please include them with your response to this question. Please also label the columns in the design matrix and model names in a way that I can understand them.

Here is a little background about the (fake) dataset. First, carefully read the documentation at the top of the input file. Note that there are 20 occasions (corresponding to April and September sampling periods in each of 10 years) and four individual covariates in the input file. The fish were tagged and recaptured during two 1-month-long (April and September) periods using standard sampling techniques. There are no missing data. Assume that all recaptures were measured without error. Here is a little more information about the covariates:

* **Sex** codes for whether the fish was a male (1), of unknown sex (0), or a female (-1). There is zero chance that they switch sexes during the study.
* **Length** is measured at capture as the total length from snout to tip of folded caudal fin in cm.
* **Mass** is measured at capture in kilograms.
* **Scaled white color** is measured from a digital photo of each fish, and is essentially a proportion of the skin (ranging from 0 to 1) that was white. It has been suggested that paler fish may be more susceptible to harvest.
* In addition, the annual harvest of this fish (in metric tons) is 300, 330, 400, 440, 600, 525, 540, 510, 575, and 560 for the 10 years. Harvest is in fall.

When you open the .dbf file in MARK, you will get the message “No output is stored in this Results Database”; click o.k. and you will be directed to an empty results browser where you can begin building models. You are free to email me if any of this is unclear. **And to keep the analysis on track, do not consider (or even run) the fully time-dependent model.**

The general approach I took to this is to first build CJS models on what my hypotheses were related to recapture probability. I did not include color as a covariate for recapture probability, based on the context of the question. Once those were narrowed down to variable(s) that were consistent in the top performing models for recapture probability, I held recapture parameters constant while changing models for Phi. For survival, I used year, season, harvest, color, gender, and length as components of the model, however, not all combinations of those factors were included. As gender became apparent as one of the most significant factors of survival, I added color and length covariates, which actually were somewhat competitive models (within 2 AICc of the top model per parameter added). Due to context of the questions and to avoid over-parameterization of the model, I only considered one model above 30 parameters, and even that made me queasy.

I expect that there were serious and significant deficiencies in my approach. Although I have much more experience in RMark than in MARK itself, I was confident at the beginning of the week that I could handle this analysis. That’s no excuse for my performance on this question, however. It really elucidates how much I rely on peer-peer learning and input from Dr. Weber on building MARK models... Before this week I don’t think I’ve run a model in MARK itself since Dr. Klaver’s class.

**After you complete your analysis, please answer the following questions:**

* 1. **What is your best estimate of the seasonal survival of fish (April to September, and September to April)? Be sure to show your work.**

I’m going to have to bite the bullet on this one and admit my failure to implement models correctly in Program MARK. My top performing model did not include an effect of season or year on survival but did include an effect of gender. The model I have with season included for survival, however, had very similar real estimates of survival (96%-98%, Figure 1), leading me to believe I’ve erred in constructing the models. My top performing model (which included gender) had survival estimates of 96.3% for males and 98.0% for females (Figure 2).

A picture containing object, clock

Description automatically generated

Figure : Seasonal survival

A screenshot of a video game

Description automatically generated

Figure : Gender Survival

* 1. **Is there any evidence that sex, length, mass, or scaled white color is an important predictor of survival or capture probability in this population? If so, then how and under what conditions? (1-2 pages)**

The way I constructed these models, sex was the biggest factor that drove changes in parameter estimates. When length and mass were evaluated as covariates for recapture probability, I found that gender still held a strong effect on recapture probability. Given the context of the problem, I did not evaluate if color had an effect on recapture probability. Length and mass were confusing covariates in this case, I assume the column values represented the mean length/mass of the fish at all captures, or perhaps just the initial length/mass at capture. Eventually, I settled on maintaining recapture probabilities dependent on season and gender.

Once the recapture probability was set, I began changing Phi to include year, season, gender, color, and length. Year and season did not have promising effects on Phi, and although I saw harvest included for Phi in the first .DBF I opened, I have to admit I failed to incorporate it effectively into these models (only one model has harvest). After that point, sex had the highest improvement in model fit above other covariates. Combining gender together with length and color at the same time resulted in a fairly competitive model compared to the model where survival was based on gender alone. Removing length from that model but keeping gender and color resulted in a model that was also competitive with the top model. Because mass and length are correlated, I chose to keep with only mass

Without gender, the survival models with just color as a covariate had the lowest model rank, and the model of survival based only on length was somewhat higher in ranking, but still outside the range of competition with top models. I would conclude from these (sketchy) results that color of the fish did not influence their catchability in the harvest, and that gender of the fish (and perhaps behavior associated with gender) had more of an effect on catchability both for sampling and for survival of harvest, with females being more likely to suffer mortality. Mass was not included in estimates of survival because, again, mass and length are highly correlated in fish and including both of them in a model would be inappropriate.

* 1. **What problems, if any, did you encounter in your analysis and how might they be resolved? (1 page)**

First off, my understanding of the Program MARK GUI is apparently at a beginner level. Following examples and working in groups to develop models is one thing but taking simulated data and contriving models from scratch is a whole other ballgame: the major leagues, even. Besides the initial confusion of opening a file already populated with models, I attempted to re-save an INP file and begin fresh in Program MARK but was pretty quickly frustrated after running an initial set of models that varied significantly from the models that showed up in the first place. I did examine their structure, but as you said those models might have introduced some bias and done more harm than good when it came to constructing my own models.

Then, I figured I would try to take this into my more comfortable realm and apply directives of this question in RMark. That idea dissipated quickly, however, as error after error prompted me to abandon that idea as well. Clearly, use of MARK and RMark is both an art and science. Returning to MARK, I focused on augmenting identity matrices and making appropriate comparisons, but I must admit that those were likely implemented incorrectly. Chapter 11 of the MARK book was helpful, but definitely too much to learn in one of the busiest weeks of my life.

Reviewing the covariates, I did observe that lengths and masses were imperfect. In some instances, fishes of a similar length were widely different in mass. This is of course common in real fisheries data but given limited time for this assignment I could not employ any systematic tests of those variables to censor data. Further, the specification for gender as non-binary 0 and 1 only probably needed to be corrected and may have introduced even more error into my analysis. I really hope that at least some parts of the approach are valid!

* 1. **Summarize all of the findings from your analysis in a short (<400 words) paragraph, just as you would write for an abstract of a manuscript.**

The effect that harvest has upon a species can have significant implications on the population’s trajectory and management needs. Species may compensate for harvest through increased natural survival, and if survival does not change between years in the presence of harvest, then that species may exhibit a completely compensatory response to harvest. Here I examine the Iowa Freshwater Drum (IFD; *Aplodinotus grunniens dinsmoreii*), an abundant species in the Midwest and an important commercial fishery. Over ten years a population of IFD was sampled during two one-month intervals in April and September, with annual commercial harvest occurring in fall/winter periods. The amount of harvest (in metric tons) was provided by commercial fishing contractors, and individual covariates of fish were recorded to explore factors that may increase or decrease an individual’s susceptibility to commercial harvest. I hypothesized that recapture probability of IFD would be affected by sample season (e.g., spring vs. fall), gender, and length. In addition, I suspected that survival would be a factor of scale color in IFD, because methods use to harvest the fish rely on visual targeting; paler fish are more likely to be seen in water. Using Cormack-Jolly-Seber models in Program MARK, I developed a suite of potential survival and recapture probability models and ranked models through AICc. The top performing models suggested that both gender and season influence recapture probability, while survival was primarily influenced by additive effects of color, gender, and length. Other individual covariates had little no effect on recapture probability, and accounting for harvest did not improve model fit. Model results indicated that harvest had little to no effect on survival, suggesting that IFD in this system are strongly influenced by density dependent processes and exhibit compensatory mortality to harvest. Therefore, relaxed restrictions on harvest of IFD in this system are not likely to induce a fishery collapse and the maximum sustained yield of IFD in this system has not been met.