FISH 622 Homework #4: Age-structured Dynamics

Curry Cunningham 2021

Instructions

Please make sure to follow the instructions in the homework and, to the extent you can, feel free to **work with others to learn from each other**. If you get hung up and can't figure out things on your own, feel free to e-mail others in this class or your peers. For homework purposes, perhaps you can work together via e-mail, Google Hangout, or whatever works best for you. However, each of you will need to submit your own completed homework assignment for evaluation.

Please post to the Blackboard Discussion Forum and feel free to e-mail me if you get hung-up, and have exhausted your immediate options (google, classmates, friends). No need to bang your head against the wall, first ask your peers for help and if all else fails feel free to email me directly.

When you do get stuck, here are a few tips:

1. If you have questions about model structure, please refer to the lecture material on GoogleDrive.
2. When fitting a model to data (whether in Excel or R), always plot out the observed and predicted values.
3. Find good starting values for Solver or the function minimizer in R by adjusting the parameter values until you achieve good visual agreement between observed data and model predictions.
4. Post a question to the class discussion board on Blackboard.

This homework assignment is due by 11:59 pm on Friday April 23, 2021.

Please submit all components of the homework assignment (i.e. word, Excel, and R script) via Blackboard, and name each file with the homework number and your first and last name (e.g. **Hwk4_FirstName_LastName.docx**, **Hwk4_FirstName_LastName.R**). If requested, please insert short answers to questions and any useful and/or necessary model output or figures within your word document.

In the event you are unable to submit your completed assignment via Blackboard, please email it to me at: cjcunningham@alaska.edu. Late homework assignments will be penalized 2 points per day overdue. Please contact me ***ahead of time*** if there are circumstances requiring late submission.

Evaluation

This homework assignment will be graded out of ***50 possible points***.

Homework Contents

- **FISH 622_Homework 4_.pdf** (this file)

- **Hw4.xlsx** Excel workbook containing the age-structured assessment model implemented in Lab #9.

Excel Age-structured Assessment Model

For this homework assignment we will work with the simple age-structured assessment model we developed in Lab #9.

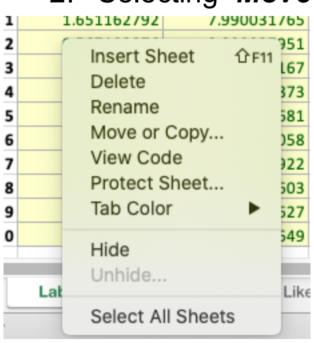
As a reminder the key attributes of our focal species *Fishitis Stickitis* stock are:

- Age-0 represents our new recruits.
- Age-12 is our plus group, which accumulates all individuals age 12+.
- Data
    - Catch age composition proportions
    - Weight-at-age is a fixed input
    - Catch in units of weight (yield) by year
    - A survey biomass index (I_t), for which the catchability coefficient (q) is directly estimated: $I_t^{pred} = q \sum_a s_a B_{t,a}$ where s_a is the selectivity at age and $B_{t,a}$ is the biomass by year and age.

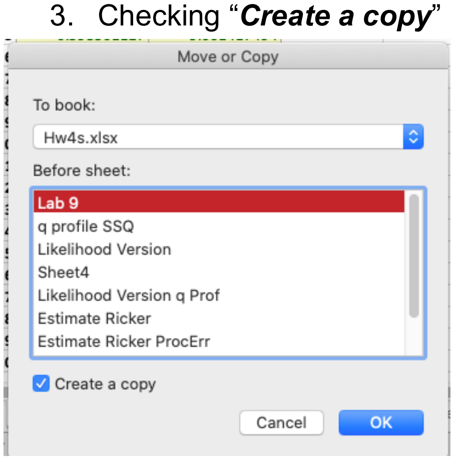
In the ***Hw4.xlsx*** file you will find my version of the model and associated figures in the tab called "***Lab 9***".

You can use this as a starting point for each problem, by:

1. Right clicking on the ***Lab 9*** tab
2. Selecting "***Move or Copy***" from the menu



3. Checking "***Create a copy***"



4. Selecting a location in which to create the copy, and clicking “***OK***”

This will allow you to create a new sheet (tab) within the Excel workbook, which is a copy of the initial model to which you can make additions or changes without altering the first sheet (tab).

Problem 1 – Conversion to a Maximum Likelihood Model (5 points)

In this problem you will make a straightforward change to how our age-structured model is fit to the available data. Specifically, you convert sum-of-squares comparisons between the natural log of predicted (I_t^{pred}) and observed (I_t^{obs}) survey indices, and the between observed (Y_t^{obs}) and predicted (Y_t^{pred}) annual fishery yield, to true likelihoods.

In Lab 9, we approximated a normal likelihood for the comparison of observed and model-predicted yields using SSQ as: $SSQ_{Catch} = \sum_t \left(Y_t^{obs} - Y_t^{pred}\right)^2$ in ***Cell C191***. Similarly, we approximated a lognormal likelihood for the comparison of observed and model-predicted survey indices using SSQ as: $SSQ_{CPUE} = \sum_t \left(ln\left(I_t^{obs}\right) - ln\left(I_t^{pred}\right)\right)^2$ in ***Cell D146***.

Please create a copy of the ***Lab 9*** tab and call it “***Likelihood***”. Here please update the SSQ comparisons for yield and survey indices to likelihoods, as follows:

Lognormal negative log likelihood for survey CPUE:

$$NLL\_CPUE = \frac{\sum_t \left(ln\left(I_t^{obs}\right) - ln\left(I_t^{pred}\right)\right)^2}{2\sigma_{CPUE}^2}$$

Normal negative log likelihood for fishery yield:

$$NLL\_Catch = \frac{\sum_t \left(\frac{Y_t^{obs} - Y_t^{pred}}{Y_t^{pred}}\right)^2}{2\sigma_{Catch}^2}$$

We will assume we know the values for the standard parameters for these likelihoods:

$$\sigma_{CPUE} = 0.2$$

$$\sigma_{Catch} = 0.05$$

Start by adding these assumed (fixed) observation error parameters to ***Cells I2 and I3***:

	A	B	C	D	E	F	G	H	I	J	K
1											
2		Data			Natural Mortality: M	0.2		SigCatch	0.05		
3		Estimated Parameters			Weighting for Agecomp (Omega)	50		SigCPUE	0.2		
4	Derived Parameters										
5						Age_0	Age_1	Age_2	Age_3	Age_4	Age_5

You can then easily reference these values in your likelihood calculations. Update your objective function to include the three likelihoods (yield, survey biomass index, and catch age composition).

Now, please call Solver to minimize your total negative log likelihood.

Problem 2 – Profile Likelihood for Survey Catchability (q) (15 points)

Now, that you have a true negative log likelihood (objective function) for your model given the data, we will explore the uncertainty in the estimated catchability coefficient (q). One way to visualize the uncertainty in the parameter estimate from a model is with a likelihood profile. To construct a likelihood profile you will:

1. Set the parameter of interest (i.e. q) to a specific fixed value.
2. Estimate all other parameters, by minimizing the negative log likelihood, with the excpetion of the parameter of interest.
3. Record the resulting negative log likelihood and the fixed value for the parameter of interest.
4. Repeat this process across a range of candidate values for the parameter of interest (i.e. q).

Please make a copy of the "***Likelihood***" sheet and name it "***Likelihood q prof***". Within this new sheet, please create a likelihood profile for the catchability parameter (q), across a range of trial values from 0.2 to 2.5, in increments of 0.1 (0.2, 0.3, 0.4 , … , 2.5). Next, calculate the difference between the negative log likelihood for each fixed value of q and the negative log likelihood when q is estimated in addition to all parameters (the maximum likelihood estimate or MLE): $\Delta NLL_q = NLL_q - NLL_{MLE}$. Plot the relationship between your trial q values and ΔNLL_q and ***include this in the Word document you submit***.

Note: The resulting figure should look like a bowl showing the relative negative log likelihood for different values of q.

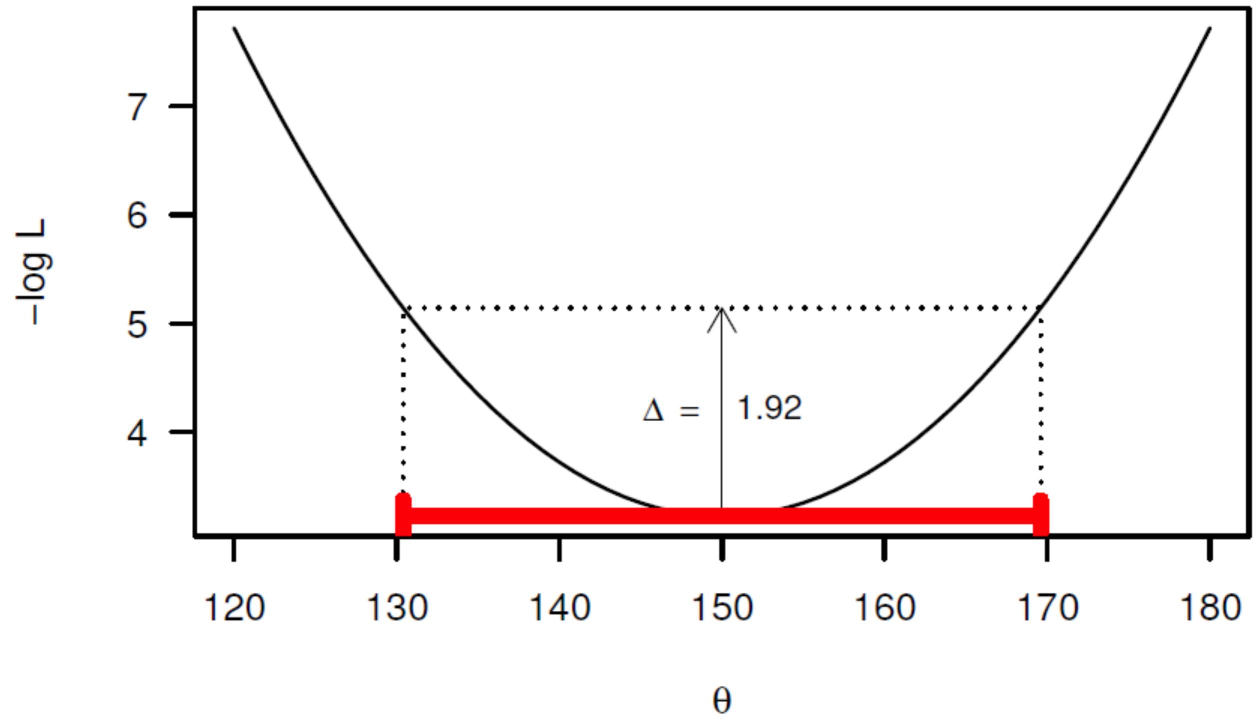
Within the word document you submit for this assignment ***describe the shape of your likelihood profile and what that implies about parameter uncertainty***. In addition, ***please describe what other parameters within your model change as you vary q*** to create your likelihood profile, and your interpretation of why these changes occur.

Problem 3 – Approximate 95% Confidence Interval (5 points)

An approximate 95% confidence interval for a parameter can be derived from its likelihood profile. The 95% confidence interval encompasses that range of parameter values that are within 1.92 log likelihood units of the MLE (i.e. the lowest negative log likelihood). For reference, the 1.92 value is derived from the Chi-square distribution with one degree of freedom: $0.5\chi^2_{df=1} = 1.92$.

Please repeat the likelihood profiling process across a narrower range of trial q values with suitably small increments of q such that you can get a good approximation of the 95% confidence interval (20 trial values for $q$should be sufficient).

Please report your approximate 95% confidence interval within the word document you submit, along with a plot of this narrower likelihood profile.



Problem 4 – Stock-recruitment Relationship (15 points)

Until this point we have assumed that annual recruitment R_t is independent of spawning stock biomass. In practice, we estimated annual recruitments within our age-structured stock assessment as independent parameters.

Let's assume that the relationship between annual spawning stock biomass and recruitment (Age-0 abundance: ***Cells C38:C57***) can be represented by a Ricker relationship. Further, we will assume that only the component of the population age-5 and older are mature and contribute to spawning stock biomass:

$$m_a = \begin{cases} 0 & a < 5 \\ 1 & a \geq 5 \end{cases}$$

Please make a copy of your "***Likelihood***" sheet and call it "***Estimate Ricker***". In this problem you will estimate the two parameters of the Ricker relationship. To do so you will need to calculate:

1. Spawning stock biomass by year (SSB_t), as the product of maturity-at-age and annual biomass-at-age ($B_{t,a}$), summed across ages
2. Use the linearized version of the Ricker relationship to predict recruitment, as: $ln(R_t) = ln(\alpha SSB_t) - \beta SSB_t$, conditional on the value of the Ricker α and β parameters.
3. Use Solver to estimate all model parameters (including q) and the two Ricker parameters.
    a. Note: For now we are ignoring process error in recruitment, and annual recruitments will no longer be treated as free parameters.

Please plot recruitment as a function of spawning stock biomass and include both this plot and the estimated Ricker parameters in your assignment word document.

Problem 5 – Stock-recruitment Relationship with Process Error (10 points)

In the context of an assessment it is likely more appropriate to incorporate process error (random variation) in the estimation of recruitment. Here you will treat the annual estimates of recruitment from the Ricker relationship ($ln(R_t) = ln(\alpha SSB_t) - \beta SSB_t$) as the expected values for the natural recruitment. Then we will estimate annual recruitments (R_t^{pred}) (***Cells C38:C57***) as free parameters as we did before, and minimize the squared differences between expected and predicted recruitments in log space. Please make a copy of the "***Estimate Ricker***" sheet and call it "***Estimate Ricker ProcErr***".

Within this new sheet you will need to:

1. Calculate the difference between the natural log of estimated annual recruitments $ln\left(R_t^{pred}\right)$ and the expected recruitment in log space (i.e. $ln(R_t) = ln(\alpha SSB_t) - \beta SSB_t$)
2. Calculate the sum of squared differences between estimated and expected log recruitments with the ***=sumsq()*** function.
3. Add the sum of squared deviations between the natural log of estimated and expected recruitment to your objective function.
4. Estimate all parameters (2 selectivity parameters, catchability, initial numbers at age, annual fishing mortalities, annual recruitments, and the 2 Ricker parameters).

In the word document for this assignment, please include the plot of the relationship between estimated annual recruitments (R_t^{pred}) and spawning stock biomass, along with the estimated values for the two Ricker parameters.

Time Allocation

At the end of the word document you submit for this homework assignment, please estimate the number of hours you spent on this assignment.