WFC 198 “Sampling Animal Populations” – Homework Week 4

Due May 2nd, 11:59pm

Please complete the following exercises in R, “knit” your script into an MS Word report and submit it on Canvas under Assignments. We will take off 10% of your points if the assignment is not submitted in the right format.

Please preface every piece of code with a comment that states to which question the code belongs, for example: “#Task 3a” before the code used to tackle Task 3a. If we cannot follow your code, we cannot grade it.

Work through the questions in order, as they build on each other.

This homework has you analyze a distance sampling data set of rabbits, sampled over 4 transects, 2 in good habitat and two in poor habitat. The data are courtesy of Dr. Mike Conroy (U of Georgia). Download the data set (“RabbitDSData.csv”) from Canvas and load the package “unmarked” (remember to comment out the “library()” command before knitting your report).

Reading the data set into R will produce a data frame with the following information:

The data has 4 rows, corresponding to the four transects. The first 28 columns (column names X1 through X28) of the data are the number of rabbit detections in 28 5-m distance intervals. Column number 29, named “hab”, holds a binary variable coding habitat quality; 1 stands for good habitat, 0 stands for poor habitat. The last column, named “tlength”, provides the length of each transect. Accounting for transects of different length (because all else being equal, there will be more rabbits on longer transects) is not something we have done before, and I will provide instructions in the tasks below for how to do that.

Task 1: Data summary statistics (3 Pts)

1. Read the data into R (1 Pt)
2. Calculate the total number of rabbit detections across all distances and transects (1 Pt)
3. Calculate the number of rabbits detected in good habitat and in poor habitat (1 Pts)

Task 2: Creating the unmarkedFrame (6 Pts)

1. For the next step (2b) you need the distance breaks, so set up a numeric vector that holds the 5-m distance breaks for the 28 distance categories, starting at 0m and going out to 140m. (2 Pts)
2. Use the unmarkedFrameDS() function to get the data into the right format for analysis with the distsamp() function. Use the R help and the R script from the Distance sampling lab for a reminder of how to specify the function arguments. Remember: this is line transect data, and distances are measured in meters. Also, we have a categorical covariate (good vs bad habitat) here, which should not be scaled. We need to specify one additional argument (compared to what we did in lab), and that is transect length. The argument to do that within the unmarkedFrameDS() function is “tlength”. Transect length has to be provided in the form of a vector; lengths of the four transects are stored in the column “tlength”, so in the unamrkedFrameDS() functions, set tlength=DataObject$tlength [note: use the name of the object holding the rabbit data in place of “DataObject”]. (4 Pts)

Task 3: Fitting distance sampling models (8 Pts)

1. Fit the following six models: (i) no covariates on detection or abundance, with a half-normal detection function; (ii) no covariates on detection or abundance, with a hazard detection function; (iii) habitat as a covariate on abundance, with a half-normal detection function; (iv) habitat as a covariate on abundance, with a hazard detection function; (v) habitat as a covariate on detection, with a half-normal detection function; (vi) habitat as a covariate on detection, with a hazard detection function (6 pts)
2. Using AIC, determine the best model out of the six models fit under a (make sure your knitted report shows the model selection table). (2 Pts)

Task 4: Results interpretation (10 Pts)

1. Produce a model summary for the best model as selected under 3b (1 Pt)
2. For that same model, and using the appropriate R command, calculate rabbit density in good and in poor habitat. Note: calculate density on the real scale, not the link scale! Make sure your knitted report shows the code as well as the numeric result of the calculation (2 Pts)
3. What is the ecological interpretation of these quantities? Write a short answer into your R script that contains the estimates (with standard errors) produced under 4b – remember to comment out anything that is not a command! (2Pts)
4. Plot the detection function for the best model over a range of distances from 0 to 140 m. Tip: We developed code to plot a detection function early on in the Distance sampling lab. In that code, you need to specify the parameters of the detection function as they were estimated by the best model. Remember to back-transform model estimates to the real scale. Specify distances at 1-m intervals. Label your axes. (5 Pts)

TOTAL POINTS: 27