This assignment has 3 questions. The first two have multiple parts.

1. 45 pts. This problem uses data from a large study of Great Cormorants (picture is here: `https://ebird.org/species/grecor`). These data are from birds banded as breeders of unknown age from 1984 to 1992 and recaptured in 1985 to 1993. The capture history data are in cormorant.txt. The m-array summary is:

| Year | # released | 1985 | 1986 | 1987 | 1988 | 1989 | 1990 | 1991 | 1992 | 1993 | never |
|------|-----------|------|------|------|------|------|------|------|------|------|-------|
| | | | | | Recaptured in year: | | | | | | |
| 1984 | 157 | 42 | 12 | 16 | 1 | 0 | 1 | 1 | 1 | 0 | 83 |
| 1985 | 174 | | 85 | 22 | 5 | 5 | 2 | 1 | 0 | 1 | 53 |
| 1986 | 298 | | | 139 | 39 | 10 | 10 | 4 | 2 | 0 | 94 |
| 1987 | 470 | | | | 175 | 60 | 22 | 8 | 4 | 2 | 199 |
| 1988 | 421 | | | | | 159 | 46 | 16 | 5 | 2 | 193 |
| 1989 | 413 | | | | | | 191 | 39 | 4 | 8 | 171 |
| 1990 | 514 | | | | | | | 188 | 19 | 23 | 284 |
| 1991 | 430 | | | | | | | | 101 | 55 | 274 |
| 1992 | 181 | | | | | | | | | 84 | 97 |

We will use the Cormack-Jolly-Seber model to estimate annual survival probabilities and capture probabilities. Note that birds may be recaptured more than once in this data set. The 174 individuals caught and released in 1985 includes the 42 that were released in 1984 and recaptured in 1985.

(a) Fit the general CJS model with year-specific survival probabilities and year-specific capture probabilities. Two parameters in this model can not be estimated. Which two parameters are those and what in the output indicates they can not be estimated?

(b) Using the model in 1a, estimate and report the survival probability for 1985-1986, and its standard error.

(c) Fit a CJS model with a year-specific survival probability and a single capture probability applied to all years. Estimate and report the survival probability for 1985-1986, and its standard error.

(d) Estimate and report the model-averaged survival probability for 1985-86 and its standard error.
Note: Remember you can specify indices= to model average only specific parameters. If one of the models you are including in the model suite has non-identified parameters, that model is dropped from the suite. So, you can't just use parameter = 'Phi'.

(e) The cormorant population size was increasing consistently over the period of these data. If there is density dependence acting on survival, you would expect $\phi$ to decline consistently from 1984 to 1992. One way to evaluate this is to fit a regression model using time as a continuous variable. I.e., $\text{logit}(\phi) = \beta_0 + \beta_1 time$. Assume a time-specific model for capture probabilities (the most general model possible). Fit this model and report the estimated values of $\beta_0$ and $\beta_1$, with their standard errors.

(f) Which of the three models (1a, 1c, or 1e) is most appropriate for these data? Briefly explain your choice.

(g) It is interesting to graphically compare survival probabilities for models 1a and 1e. Plot survival probability (Y-axis) over time (X-axis) from model 1a, with 95% confidence

intervals and overlay the fitted line from model 1e. Do you have any concerns about lack of fit of model 1e? Your answer is the plot and list of any concerns, with short explanations.

2. The data in huggins.txt on the class web site are individual capture histories for a fictitious animal in a closed population three occasion sampling scheme. Various covariates were recorded for each animal: the size of its home range, its sex (0=male, 1=female), weight when first captured, and the type of trap (3 types, so 2 columns of 0/1 values in the data). Size of home range and weight have been standardized to mean 0, variance 1 to improve numerical performance. For most parts of this question, we will ignore behavioral and time effects on the capture probability.

   (a) Consider 4 covariate models: intercept + home range, intercept + sex, intercept + weight, and intercept + trap type (3 levels). Which variable appears to have the most influence on capture probability? Explain your choice.
   Note: You don't need to consider models with more than one covariate. If you want to use an automated procedure to explore many possible covariate models, the dredge() function in the MuMIn (multi-model inference) package can process RMark models. dredge(), for those not familiar with it, will troll through many models and summarize the results.

   (b) Using the model chosen in the previous part, estimate the population size, N.

   (c) You wonder what happens if you don't incorporate heterogeneity. You decide that model M0 is the appropriate "no-heterogeneity" model. What is the estimated population size under M0?

   (d) The log likelihood for the M0 model, using my log likelihood function is -762.97. Is it appropriate to compare this lnL or associated AIC to AICs from the models in question 2a. Explain why or why not.

   (e) Fit two 2 component mixture Pledger models. Both will have no behavioural effects, so capture and recapture probabilities are the same values:
       model A with the same probabilities for each capture occation (no time effects).
       model B that allows capture probability to vary by time, but some individuals are still "hard to capture" and others are "easy".
   Does either model suggest evidence of heterogeneity in capture probabilities? Explain why or why not.

   (f) Look at the results of fitting the first 2 component mixture Pledger (model A). Is there anything in these results that suggests a 2 component mixture is not a reasonable model? Briefly explain your answer.

3. Work in pairs on this problem. Each pair will find a paper that uses a open population model (or perhaps multiple). Prepare a short (7 minute) presentation describing:

   • The biological background of the study

   • Its goal, e.g., what did the authors want to estimate or test?

   • The model they adopted, in enough detail to give your audience a sense of how the model was set up and used for the problem.

   • Briefly, the results.

The purpose of these presentations is to share the diversity of models and applications of open population models. You are not expected to do any computation, even if the raw data are provided as part of the paper.

For example, for the multi-site bird study described in lecture on 12 October, a summary of a presentation I would give would be:

- A large population of a rare bird in Wales is declining, which raises conservation concerns. There are three much smaller populations in the vicinity of the large population that are previously unstudied. There have been various schemes proposed to increase the population growth rate in the large population.

- Estimate population growth rates and migration rates among the four populations.

- (This is a guess): A multistate CJS model to estimate migration probabilities among populations combined with Pradel reverse time models to estimate population growth rates for each population.

- Population growth rates were much larger in the 3 smaller populations. A large proportion of the birds from the small populations migrate to the large population. It seems that the large population is a sink, maintained by immigration from the smaller populations. This suggests conservation concern should be directed to minimizing disturbance or other impacts to the smaller populations.