1. These data come from a study of a deer mice population at one site in Pennsylvania. Individuals were trapped on 5 nights. Summary statistics are:

<div align="center">

Trap night

|	1	2	3	4	5	
n_i	20	14	12	11	11	
u_i	20	10	11	7	6	
m_i	0	4	1	4	5	
M_i	0	20	30	41	48	$M_{t+1} = 54$

</div>

The variables are named the same way they are named in the lecture notes; these are the usual names found across the literature. Because the data are only available in this summarized form, you will need to use optim() and log likelihood functions to estimate parameters.

Log likelihoods for models M_0 and M_t are given in the lnlM0t.r file, although note that I omit the optional constants from the log likelihood calculation. The sufficient statistics used in my functions (and in the equations in my notes) are:

t	Number of trapping occasions
n	number caught on each occasion (vector of length t)
$n_.$	ndot: total number of captures, $n_. = \Sigma_{i=1}^{t} n_i$
M	number of tags in population just prior to occasion i, $M_i = \Sigma_{j=1}^{i-1} u_j$ except $M_1 = 0$.
$M_.$	Mdot $= \Sigma_{i=1}^{t} M_i$, note: does not include M_{t+1}.
M_{t+1}	Mt1: Total number of unique individuals seen
u	number of unmarked individuals caught on each occasion (vector of length t)
m	number of marked individuals caught on each occasion (vector of length t)
$m_.$	mdot: total captures of marked individuals, $m_. = \Sigma_{i=1}^{t} m_i$
	note: $n_i = u_i + m_i$ for all i.

(a) Estimate and report the population size, N, and its standard error under model M_0.

(b) Estimate and report the population size, N, and its standard error under model M_t.

(c) Estimate and report the population size, N, and its standard error under model M_b. Note: You will have to code the log likelihood function for model M_b. If you want to compare lnL or AIC with the results from "my" functions in lnlM0t.r, you need to omit the constants.

(d) Which model, M_0, M_t, or M_b is the best for these data? Briefly explain your choice.

(e) If M_0 was your answer to question 1d, would it be appropriate to conclude something like "model M_0 is a good fit to the data."? Briefly explain why or why not.

(f) Using AIC weights, compute the model averaged estimate of N and its standard error, using both the Buckland se formula and the Revised formula.

(g) The log likelihoods I provided for M_0 and M_t use N as the parameter. You will sometimes see f_0 used instead of N. f_0 is the number of "never seen" individuals, i.e., n_{000} for a 3 capture occasion study. Note that $N = M_{t+1} + f_0$. What is the equation for the log-likelihood of model M_b when f_0 is the parameter instead of N, i.e. what is $\ln L(f_0, p|n_., M_{t+1})$?

(h) Consider separately fitting both a model using N and the same model using f_0 and calculating the lnL at the appropriate mle's for each model, i.e., at either $(\hat{N}, \hat{p})$ or $(\hat{f}_0, \hat{p})$ if you were fitting M_0. Would you expect to get the same value of the log likelihood for these two models, when each lnL is evaluated at the appropriate mle's? Briefly explain why or why not.

(i) Using results from fitting model M_b, calculate a Wald 95% confidence interval for N.

(j) Using results from fitting the model M_b in terms of $\log f_0$, use log transformed f_0 to calculate a Wald 95% confidence interval for N. In other words, use $\log f_0$ as the parameter, estimate the se of $\widehat{\log f_0}$, calculate the confidence interval for $\log f_0$ and use that to compute a confidence interval for N. In case you care about this detail, M_{t+1} should be treated as a constant when computing the confidence interval for N.

(k) Using model M_b, plot the profile likelihood for a range of values of N around the mle, $\hat{N}$.
Note: For model Mb, the conditional mle's of p and c given $\hat{N}$ are $\hat{p} = \frac{M_{t+1}}{t\,\hat{N} - M_.}$ and $\hat{c} = \frac{m_.}{M_.}$.
Your answer is the plot.

(l) Use the plot, and possibly additional numerical investigations, to calculate a 95% profile confidence interval for N when estimated with model M_b.

(m) Repeat question 1k plotting the profile likelihood as a function of $\log N$ and again as a function of $\log f_0$. Your answer is the two plots.

(n) Look at the plots of profile likelihood vs. N, profile likelihood vs. $\log N$, and profile likelihood vs. $\log f_0$. Use these plots to decide the best quadratic approximation: that is which is more appropriate, to use a use a quadratic approximation in terms of N, a quadratic approximation in terms of $\log N$, or a quadratic approximation in terms of f_0? Briefly explain your choice.
Notes: It may help to refer to one or more plots as part of your explanation. You can use the addQuad() function to compute (and hence plot) the quadratic approximation given results (estimate, lnl, and se of the estimate) of a fit.

(o) Wald intervals are much more commonly used than profile intervals (at least in most of today's computing packages) because they are much simpler to compute. Based on your results from questions 1k, 1m, and 1n which would you recommend: Wald intervals based on N, based on $\log N$, or based on $\log f_0$?. Briefly explain your answer.

2. You have been studying a population of dusky-footed snouters, one of many species of mammal in the order Rhinogradentia known to occur only on Hiddudify, an island in the Hy-yi-yi archipelago in the Southern Pacific Ocean. (Stümpke, H. 1967. *The Snouters: Form and Life of the Rhinogrades.* U. Chicago Press. see also Wikipedia/ Rhinogrades or `https://spec-evo.fandom.com/wiki/The_Snouters:_Form_and_Life_of_the_Rhinogrades`.
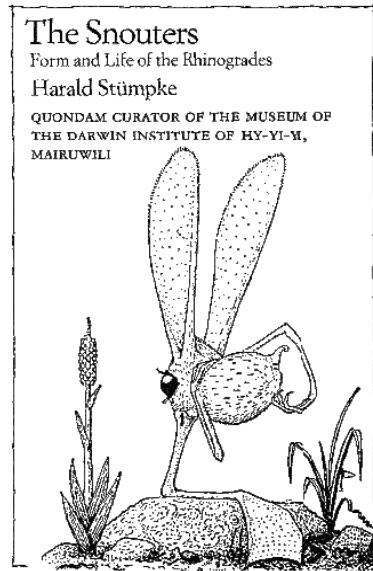
Dusky-footed snouters are quite rare, but they are not yet listed as an endangered species. You trapped and marked individuals six years ago and returned last year to trap and mark individuals again. Dusky-footed snouters are short-lived; none of the individuals marked 6 years ago were seen last year. After you finished trapping last year, you tell a friend that something interesting is going on: you caught 273 different dusky-footed snouters 6 years ago, but only 142 different ones last year! Your friend is immediately concerned. He thinks the population size is obviously dropping rapidly, and we need to do something immediately.

The capture histories for 6 years ago are in snouter1.txt; those for last year are in snouter2.txt. The populations should be assumed closed within the capture period each year.

Analyze the data and figure out what is happening to the population. Is or is there not cause for concern?

Your answer should be in the form of parts of a scientific paper. You should include:

- A statistical methods paragraph (or perhaps paragraphs) that outlines the methods you used. This should be in sufficient detail for a reader to repeat your analysis. For example: "I used likelihood to estimate parameters." is too brief. "I read in the data file using import.chdata(), created these (list of formulae) models, then $\cdots$, then $\cdots$" is far too detailed. If you used Mark or RMark, just give what a reader needs to repeat the analysis; you don't need to describe the inner workings of Mark. If you wrote R code, e.g., log likelihood functions or RMark calls, those should be included in an appendix.
  Think about what you would need if you wanted to reproduce an analysis. That's what this paragraph (or few paragraphs) needs to include.

- A results section with appropriate tables and / or figures.

- A one paragraph conclusions section.

Pay attention to what is **not on that list**:
- no introduction section,
- no non-statistical methods section,
- no discussion section.