Homework # 1, September 21, 2009

Due: 4 pm, 6 Oct 2009, to my office (2121 Snedecor) or mailbox (in 1121 Snedecor)

Reminders:

1. Choose **4 of the 5** problems. You are not expected to do all five.
2. Please see the **Homework guidelines** page of the class web site for homework policies and the format for answering the data questions (1 and 3).
3. In particular, you are encouraged to work together. However, please write up your own answers.
4. This is not a progamming class. I will not provide functions (except for some of problem 4), but I will certainly help you get your functions working.

Problem 1. The following data come from a study of cottontail rabbits, *Sylvilagus floridanus*. Animals were captured for 7 successive nights. Each morning, captured animals were uniquely marked and released. Because there are so many different capture histories, I'm giving you the data in summarized form. Remember, the number marked in the population (M_i) is the number of marked animals prior to that capture occasion.

Capture occasion	1	2	3	4	5	6	7	
# caught, n_i	9	8	9	14	8	5	18	
# newly caught, u_i	9	6	3	11	4	1	10	
# marked, m_i	0	2	6	3	4	4	8	
# marked in pop, M_i	0	9	15	18	29	33	34	44

You may assume a closed population. The investigators want to estimate the number of cottontails in this population. The investigators have just started working with this species, so they do not know which model (M0, Mt, Mb) is most appropriate. What can be said about the population size?

Note: the sufficient statistics for each model can be computed easily from this summary table. You do not need to know the capture histories. For example, for model M0, the total number of captures $= \sum n_i = 71$.

Problem 2. Dr. Otis briefly mentioned model M_{tb}. He also mentioned that this model has a problem: there are more parameters than there are sufficient statistics. This is a model that allows capture probability to vary among capture occasions and also with behaviour (i.e. between 1st and subsequent captures). For capture occasion i, an individual has P[capture] = p_i] if that animal has not previously been caught and P[capture] = c_i if it has. Consider three capture occasions (i.e. a total of 8 possible capture histories).

a) How many parameters are in Model M_{tb} for 3 capture occasions?
b) Write out the 8 possible capture histories and their associated probabilities.
c) Write out the log-likelihood function for a multinomial model using 8 possible capture histories. Combine terms wherever possible.
d) Use the simplified expression of the log-likelihood to determine the sufficient statistics. How many sufficient statistics are there?
Note: remember that linear combinations of sufficient statistics do not count as additional sufficient statistics. I.e., if A and B are sufficient statistics, A-B is not a third sufficient statistic.

Problem 3. This problem explores a model that allows some between-individual heterogeneity in capture probability. Males and females may have different capture probabilities, but there is no heterogeneity within males or within females. For many species, it is possible to determine the sex without error when you capture them. Hence, your data include the number of males with a specific capture history and the number of females with a specific capture history.

Consider a sex-specific generalization of model M_t with 3 capture occasions. You are interested in estimating the total population size $N = N_m + N_f$. You will also need to estimate 6 other parameters, p_{m1}, p_{m2}, p_{m3}, p_{f1}, p_{f2}, and p_{f3}.

a) Derive the log likelihood function to estimate N and the other parameters.
b) What are the sufficient statistics in this model? Please simplify them as much as possible. Is it possible to estimate all the parameters?
c) Could you estimate N by separately estimating N_m using only data from the male individuals and N_f using only data from the female individuals, then adding together $N_m + N_f$?
Hint: If you wrote down a log likelihood using N_m and N_f instead of N, is it possible to separate the log likelihood into a one part that involves only 'male' parameters and data and a second part that involves only 'female' parameters and data?

Problem 3. One of the classic data sets in the mark-recapture world was collected by French scientists on the European Dipper. This bird lives alongside mountain streams. Although data were collected annually from 1981-1987, we'll consider just two of those years (1984 and 1985). We'll also consider the population closed over those two years. The data are in dipper.txt. The first two columns are the capture history (0 = not captured; 1 = captured). The next two fields are the number of males and number of females with that capture history.

The investigators are interested in:
a) The number of dippers in the study area.
b) Whether capture probabilities for males are the same as those for females.

What can you tell them?

Problem 5. The following data were concocted to make a point about estimating trends in abundance. For the last couple of years, you have been studying a population of dusky-footed snouters, one of many species of mammal in the order Rhinogradentia only known to occur on the island of Hiddudify (Stümpke, H. 1967. *The Snouters: Form and Life of the Rhinogrades.* U. Chicago Press. see also Wikipedia/ Rhinogrades). Dusky-footed snouters are quite rare, but they are not yet listed as an endangered species. Each year, you ran a mark-recapture study with 4 trap nights to estimate the population size. After you finished trapping in the second year, you tell a friend that something interesting is going on: you caught 216 different dusky-footed snouters in the first year, but only 111 different ones this year! Your friend is immediately concerned. He thinks the population size is obviously dropping rapidly, and we need to do something immediately.

Here are the data:

Year 1:

Capture occasion	1	2	3	4	
# caught, n_i	118	70	56	153	
# newly caught, u_i	118	36	24	38	
# marked, m_i	0	34	32	115	
# marked in pop, M_i	0	118	154	178	216

Year 2:

Capture occasion	1	2	3	4	
# caught, n_i	48	17	14	46	
# newly caught, u_i	48	17	10	34	
# marked, m_i	0	0	4	12	
# marked in pop, M_i	0	48	65	75	109

The population can be considered closed within each year (i.e. across the 4 trap nights each year). What do you conclude about the population sizes (last year, this year) and the change in population size? Is there anything interesting going on?