BIOL 133 – Ecological Models and Data – TEST #1 – 3 March, 2018

NB: Each subproblem is worth 10 points. There are a total of 160 points.

NB: There is a table of functions and formulas at the end of the test (pg. 6)

NAME: \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

1.  A population of Meleager’s blue butterfly, *Polyommatus daphnis*, was surveyed in 1993 near the town of Piatigorsk, which was the source of uranium for the first atomic bombs made in the Soviet Union. During the survey researchers captured 150 butterflies, of which 30 were male, 48 were female, and 72 were gynandromorphs with both male and female features (see Dantchenko et al. 1995 *Holarctic Lepidoptera* 2:77-79).
   1. Calculate , the maximum likelihood estimate (MLE) of the proportion of gynandromorphic butterflies.
   2. Six specimens of this butterfly were collected from this region prior to 1940, i.e., prior to development of radioactive industry in Piatigorsk. Using your estimate of from 1a, calculate the probability that 0/6 butterflies were gynandromorphs due to demographic stochasticity alone, as opposed to a change in the frequency of this mutation.
   3. Calculate the probability that 0/6 butterflies are gynandromorphs assuming the proportion of gynandromorphs is 0.01, a typical “background” frequency in this butterfly species.
   4. Use Bayes’ Theorem to calculate the posterior probability of the two models (i.e., two hypothesized estimates of historic gynandromorph frequency) evaluated in 1b and 1c, given the data from prior to 1940 (six butterflies, none of which were gynandromorphs).
2. During the winters of 1999 and 2000, parasite loads of wild canids were monitored in the Tvier and Smoliensk regions of western Russia (Pozio et al. 2001 *Journal of Parasitology* 87:1498-1501). The table below shows the proportion of wolves, foxes and raccoon dogs which carried *Trichinella*, the parasitic roundworms that lead to trichinosis.

|  |  |  |
| --- | --- | --- |
| Species | # caught | # infected |
| Wolf | 82 | 80 |
| Red fox | 29 | 14 |
| Raccoon dog | 5 | 1 |

* 1. Calculate the MLE proportion of infected animals and 95% confidence limits for each canid species.
  2. Use a binomial glm to calculate the log-likelihood of a model with a single rate of parasitism for all three species. Report the line of R code used to run the model, as well as the log-likelihood and number of parameters.
  3. Use a binomial glm to calculate the log-likelihood of a model with a different rate of parasitism for each of the three species. Report the line of R code used to run the model, as well as the log-likelihood and number of parameters.
  4. Use a likelihood ratio test to decide which model is better for these data. Report the p-value, test degrees of freedom, and your conclusion (in words).

|  |  |
| --- | --- |
| Species | % parasitized |
| Wolf | 33.3-65.2% |
| Red fox | 6.8-37.5% |
| Raccoon dog | 22.2-39.9% |

* 1. Earlier studies, from 1979-1993, reported the rates of *Trichinella* parasitism shown in the table to the right. Do you think the proportion of parasitized animals has changed for each species? Briefly justify a conclusion for each, drawing on the analyses you conducted above.



1. Starting in the late 1970’s, the distribution of the (European) common frog, *Rana temporaria*, was monitored in wetlands in the city of Yekaterinburg, Russia (Vershinin et al.

2015 *Scientific Data* 2:150018). A portion of their data are shown in the table below:

|  |  |  |
| --- | --- | --- |
| Survey year | # wetlands searched | # wetlands occupied |
| 1980 | 10 | 6 |
| 1990 | 11 | 5 |
| 2000 | 11 | 4 |
| 2010 | 9 | 2 |

* 1. Fit a model to the data in which the number of occupied wetlands is constant throughout the four decades of the study. Report the line of R code used to construct the model, the estimate of the proportion of occupied wetlands, the number of parameters in the model, and the log-likelihood.
  2. Fit a model to the data in which the number of occupied wetlands is shows a trend over time during the four decades of the study. Report the line of R code used to construct the model, the estimate of the proportion of occupied wetlands, the number of parameters in the model, and the log-likelihood.
  3. Select the best of these two models for this data using AIC. Show the AIC for each model, and indicate which is better, based on the data.
  4. Calculate relative support for each model using Akaike weights.
  5. Do you believe variation in wetland occupancy is changing through time or is varying only due to sampling error? Justify your conclusion based on the analyses in parts a-d (I am not asking you to conduct new analyses in this subproblem).

|  |  |  |
| --- | --- | --- |
|  | Lead level | |
| Species | Normal | elevated |
| Common eider | 50 | 1 |
| King eider | 51 | 3 |
| Spectacled eider | 34 | 4 |
| Steller’s eider | 9 | 1 |

1. Eiders are sea ducks that breed in the northern hemisphere. Many – possibly all – species of eider are in decline, possibly due to environmental pollution. The data to the right are from a study of lead concentrations in tissues of four eider species collected from the Arctic Ocean (Stout et al. 2002 *Environmental Pollution* 119:215-226).

Consider two hypotheses about variation in the proportion of eiders with elevated lead in their tissues: (A) Species differ in their exposure to lead, (B) Observed variation is due to binomial sampling variation only.

* 1. Briefly (in 1-3 sentences) outline a statistical analysis that you will use to evaluate these two hypotheses. (NOTE WELL: There is more than one right way to conduct this test. We have learned at least 3 ways to do this in class. Any one of these is acceptable!)
  2. Conduct the test outlined in part a. Show as much work as needed for me to follow your steps. Report your conclusions about the two competing hypotheses.



Possibly-useful functions and formulas:

|  |  |
| --- | --- |
|  | dbinom(k,N,p, log = TRUE/FALSE) |
|  | logLik(model.name) |
| glm(Y ~ 1, family = binomial(link = “identity”)) |
|  | glm(Y ~ -1 + treatment, family = binomial(link = “identity”)) |
|  | glm(Y ~ year, family = binomial(link = “logit”)) |
|  | glm(Y ~ -1 + treatment, family = binomial(link = “logit”)) |
| AICi = -2log(Li) + 2k1 | AIC(model.name) |
| library(bbmle); library(lmtest) | plogis(variable.name) |
| lrtest(model1, model2) | ICtab(model1, model2, … weights = T) |
| cbind(successes, failures) | confint(model.name) |