BIOL 133 – Ecological Models and Data – TEST #1 – 27 February 2020

FYI: There is a table of functions and formulas at the end of the test (pg. 5)

Round **final** answers to 3 or 4 decimal places.

NAME: ____

1. In 1938, Herbert Friedman and Malcom Davis wondered if parrots tended to be right or left "handed". That is, do they tend to prefer using their right or left foot to manipulate food, or do they have no preference? They gave parrots apple slices and noted which foot they used to pick them up. Each bird was tested 20 times. They published their data in *The Auk*.

Name*	**left**	**right**
Charlie	14	6
Lara	13	7
Iago	12	8

Data in parrots.csv

*Parrots and data are real; names have been changed to protect the innocent.

a. (5 pts) If parrots don't have a preference for which foot to use, then the probability of using their left foot should be 0.5. For each parrot, use the binomial distribution to calculate the probability of seeing the observed data (number of left foot uses out of 20 tests) given a model of no preference.

b. (3 pts) Given a model of no preference (p = 0.5), calculate a likelihood for the entire dataset

c. (3 pts) Calculate the MLE proportion of left foot uses for the entire dataset. Repeat your likelihood calculations for this model (p = MLE p)

d. (8 pts) Use Bayes' Theorem to calculate the probability of the models in part b and c, given the data, assuming that these are the only two models and using uninformative priors.

e. (8 pts) In 2011, Brown and Magat published a study on 322 parrots in *Biology Letters* where 47% of the parrots were left-handed. Use this information to inform prior probabilities for yoru two models. That is, set the prior probability for the second model (p = MLE p) to 0.47. Again, assume that "left-handed" (p = MLE p) and "no preference" (p = 0.5) are the only possible models (We're ignoring the possibility of right-handed parrots for the sake of simplicity!). Re-calculate the probability of each model, given the 1938 data, using these informative priors.

2. In homework 4 you were introduced to the ball gall fly, *Eurosta solidaginis*. The ball gall fly is parasitized by a wasp, *Eurytoma gigantea*, which uses its ovipositor (egg-laying organ) to drill through the gall and lay an egg inside the developing fly larva. One of the hypothesized benefits of forming galls is protection from parasitoids like *Eurytoma gigantea*. Researchers tested this hypothesis by presenting galls of different sizes to *Eurytoma gigantea*. For each gall, they recorded the gall size, whether the wasp attempted to oviposit, and then dissected the gall to see if she succeeded in laying an egg.

Data in eurytoma.csv

Gall Size (mm)	Successes	Failures
<12.5	13	2
13.0-17.5	9	3
>18.0	1	15

From Weis, Abrahamson, and McCrea 1985

A *Eurytoma* female unfolding her ovipositor. Photo by Vida van der Walt

a. (3 pts) Find the MLE of the proportion of successful oviposition attempts by *Eurytoma gigantea* using all gall size classes. Report this proportion:

b. (3 pts) Find the MLE of the proportion of successful oviposition attempts for each gall size class separately

c. (10 pts) Use a likelihood ratio test to compare these two models (one with a single oviposition success rate and one in which rates differ by gall size class). Report the log-likelihoods of each model, and the χ^2 statistic, degrees of freedom, and p-value from the LRT. State your conclusion.

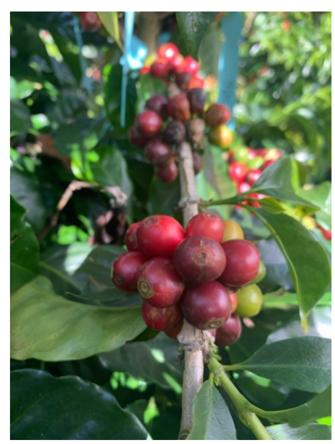
d. (5 pts) Calculate and report a 95% confidence interval for the oviposition success rate on each gall size class. Using these confidence intervals, determine which size class differs significantly in oviposition success from the others. Explain your reasoning.

3. Gabriela Garcia, a Tufts graduate student, and Madeline Bondy, a Tufts undergraduate, study coffee production in Costa Rica. Coffee plants produce fruits on the ends of branches (see photo). Branch dieback can limit the amount of space for the production of new fruits. One possible reason for this is coffee plants may deplete their resources making fruits which results in fewer resources to maintain vegetative growth and more dieback after fruiting. Gabriela estimated the number of fruits on coffee plants in June and tracked how many branches lived or died the following year across several farms.

Data in coffee.csv

Farmer	Mean # Fruits in June	# Dead branches	Total Branches
Fernando	91.4	10	80
Fernando	86.2	15	78
Fernando	88.4	11	80
Fernando	81.7	4	80
Miguel	49.3	19	87
Miguel	48.6	17	86
Miguel	52.8	6	87
Miguel	51.8	10	86
Oscar	73.8	14	71
Oscar	68.5	19	72
Oscar	76.9	21	72
Oscar	70.4	19	72

a. (5 pts) Use a likelihood ratio test to test if the proportion of dieback has a linear relationship to average fruit production in the previous June. Report LRT results and justify your conclusion.

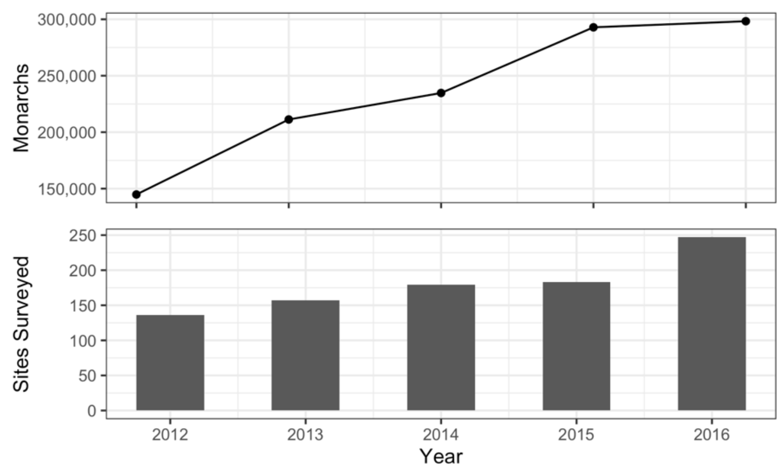


b. (5 pts) An alternate hypothesis to explain these data is that branch dieback depends mostly on farmer practices. Use AIC to compare a null model (one dieback proportion), a linear relationship between dieback and fruit production, and different dieback proportions for each farmer. Report the AIC for each model, state and justify your conclusions.

c. (3 pts) If these coffee farms/farmers were randomly sampled from all the coffee farms in Costa Rica, and the coffee plants were randomly sampled within each farm, what would the scope of inference be for this experiment?

4. West of the Rocky Mountains, Monarch butterflies migrate to the California coast to spend the winter. The Xerces Society leads a citizen science effort to monitor the abundance of Monarchs in the west by having volunteers count Monarchs at overwintering sites over Thanksgiving weekend every year since 1997. The population declined from 1.2 million in 1997 to only about 59,000 in 2009. From 2010 to 2016 the population appears to have increased, however the number of sites monitored also increased.

Data in monarchs.csv



a. (8 pts) Write the glm() code you would use to model a change in the number of monarch butterflies per site over time. Use an offset to take into account differences in sampling effort for each year.

M1 <- glm(

b. (3 pts) Use the coefficients from the model in part a to write the equation of the line describing the relationship between log(monarchs/site) and year.

c. (3 pts) Use this equation to predict the number of monarchs per site in 2012.

Possibly useful functions and formulas:

$P(x = k \mid N, p) = \binom{N}{k} p^k (1-p)^{N-k}$	dbinom(k,N,p, log = TRUE/FALSE)
$\hat{p} = \frac{k}{N}$	logLik(model.name)
	glm(Y ~ 1, family = binomial(link = “identity”))
$\widehat{\sigma^2} = Np(1-p)$	glm(Y ~ -1 + treatment, family = binomial(link = “identity”))
$P_A \approx \frac{N_A}{N}$	confint(model.name)
$P(model \mid data)\frac{P(data \mid model)P(model)}{P(data)}$	cbind(successes, failures)
	AIC(model.name)
$AIC_i = -2\log(L_i) + 2k_1$	lrtest(model.name1, model.name2)
plogis(), exp()	library(lmtest)