Exercise 8

ST2304 Exercises Week 13: Binomial Models

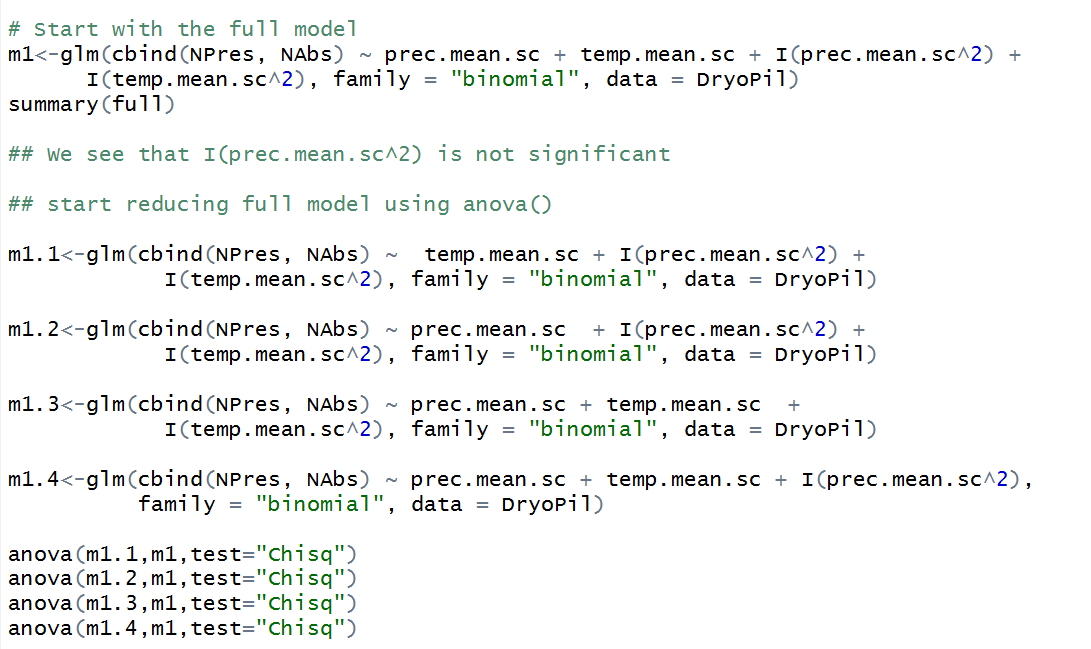
**Solutions**

**Question 1**

Fit the model with linear and quadratic terms, as above. Use anova() to compare the models (or use AIC if you want to try that!)

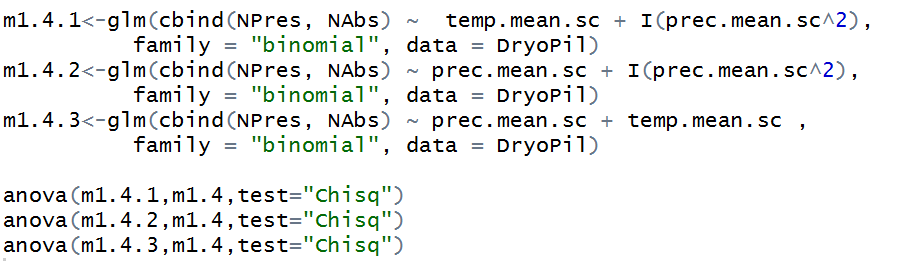
1. Would you include the precipitation and temperature in the model?
2. Would you include quadratic terms for either or both of these models?

We will do backwards model selection starting with a full model with precipitation and temperature and their quadratic terms:

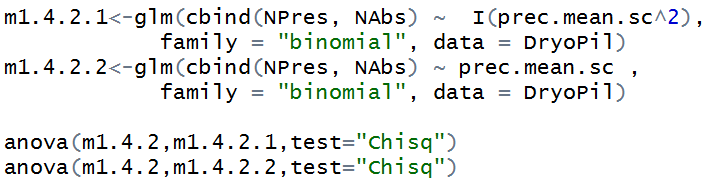


The output shows a significant p-value for all model comparisons except between models m1 and m1.4. So we use model m1.4 as a reduced form of the full model and continue reducing it, one term at a time.

The summary of m1.4 shows that the variable “temp.mean.sc” is not significant.



The output shows a significant p-value for all model comparisons except between m1.4 and m1.4.2. So we continue reducing model m1.4.2, one term at a time.

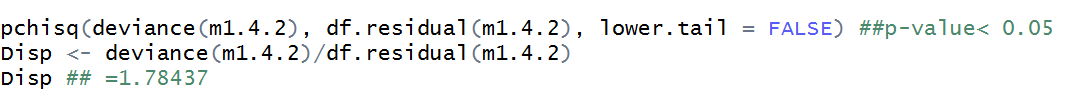


Model m1.4.2 cannot be reduced any further and is our best and final model. Thus, we include only precipitation and its quadratic term in the model.

1. Look at the parameter estimates, using summary(). Qualitatively, are there any differences between the models? Obviously the parameters will be different, but are there any large differences, and can you explain why? (if you can’t work this out from just the summary, try plotting the predictions).

Comparing summaries of models m1, m1.4 and m1.4.2, we see that the parameter estimates do not change drastically, probably because the variables that have a significant effect (prec.mean.sc + I(prec.mean.sc^2)) are included in these models.

1. Is there any evidence of over-dispersion in the model? If there is, does it change the model you would chose, and if so, how?

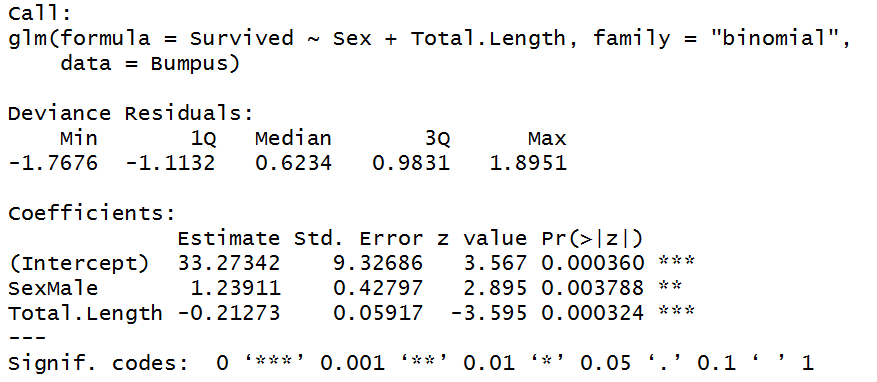


The low p-value in the goodness of fit test suggests that the data does not fit the model well. This is because of overdispersion in the model (estimate of overdispersion is 1.78). What we can do is to correct for this overdispersion by increasing the standard errors by √1.78 times.

**Problem 2: Bumpus’ Sparrows**

Fit a model with Total Length (“Total.Length”) and Sex as covariates (just as main effects), and survival as a response.

1. What effect does the total length have: do larger or smaller birds survive more?



Increase in length decreases the probability of survival. So smaller birds survive more/better.

1. What is the difference in the odds of survival of birds that are 5mm different in length?



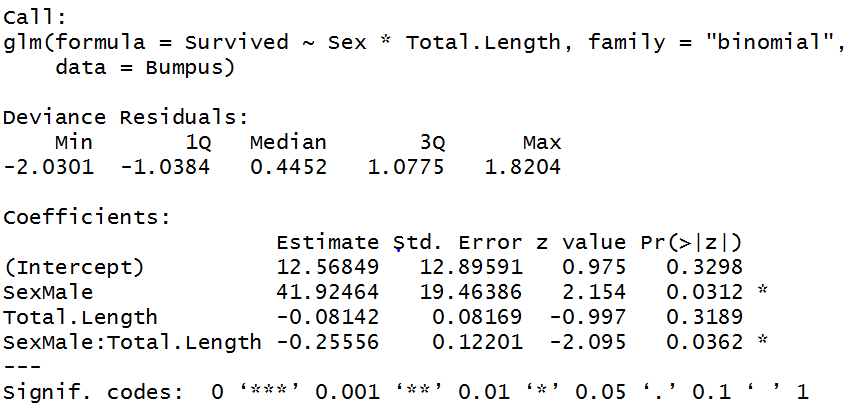
For birds with a length difference of 5mm, the difference in odds for survival is 0.34. So a bird that is 5 mm smaller than another bird has 34% higher odds of survival.

1. Does Sex have an effect on survival?

Sex has a significant effect on survival, with males surviving better than females.

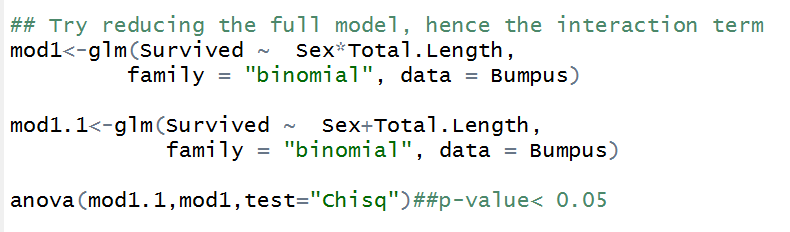
Add an interaction between Sex and Total Length.

1. What is the effect of the interaction? Is the estimated effect of body length stronger or weaker in males (i.e. is the slope of the effect steeper)?



The effect of total length is even more negative for males (goes from an estimated effect of -0.08 to -0.25). Thus, the negative effect of total length on survival is even stronger for males, compared to females.

1. Use analysis of deviance (i.e. anova()) to compare the models with and without an interaction. Would you include an interaction?



A significant p-value means that the interaction term is significant and cannot be reduced. We therefore need to keep it in the model.