STAT 601_002: Statistical Methods I, Fall 2019
24hour Take-Home Exam due on 5:00 pm, Nov. 24

Problem:

Bumpus Natural Selection Data. Hermon Bumpus analyzed various characteristics of some house sparrows that were found on the ground after a severe winter storm in 1898. Some of the sparrows survived and some perished. The data on male sparrows are survival status (1 = survived, 2 = perished), age, the length from tip of beak to tip of tail (in mm), the alar extent (length from tip to tip of the extended wings, in mm), the weight in grams, the length of the head in mm, the length of the humerus (arm bone, in inches), the length of the femur (thigh bones, in inches), the length of the tibio-tarsus (leg bone, in inches), the breadth of the skull in inches, and the length of the sternum in inches.

Data and variables: 51 male sparrows that survived (SV = 1) and 36 that perished (SV = 2) during a severe winter storm: Age (AG) is 1 for adults, 2 for juveniles; TL is total length; AE is alar extent; WT is weight; BH is length of beak and head; HL is length of humerus; FL is length of femur; TT is length of tibio-tarsus; SK is width of skull; and KL is length of keel of sternum.

Analyze the data to see whether the probability of survival is associated with physical characteristics of the birds. This would be consistent, according to Bumpus, with the theory of natural selection: those that survived did so because of some superior physical traits.

As a statistician, you propose two best models and describe why these two models are the best to explain the relationship between the probability of survival and physical characteristics of the birds. Make sure to interpret the estimated coefficients. This should be done using all the available data. Moreover, we are interested in comparing the two proposed models in term of predictivability. Use the cross validation to compare the predictivability of the two models.

Instruction:

- You can communicate about this problem with your instructor only. Don't discuss about this problem with other students except your partner.
- Both you and your partner will submit ONE report in addition to the R-code for your analysis.
- You should make written report including introduction, model/methods, result, conclusion/discussion, and appendix.
    - ✓ Introduction: you explain about the summary of data and give the goal of data analysis
    - ✓ Model/Methods: you explain how you find your best two models
    - ✓ Result: you summarize your results. If you need, you can use table or figure
    - ✓ Conclusion/discussion: you summarize and discuss about your finding.
    - ✓ Appendix: you attach your code in appendix.

Your written report should be submitted on the deadline.