Statistics 601, Spring 2026
Assignment 2 – Part 1

There are a number of issues involved in the problem of developing a hierarchical model for relating length and weight in large mouth bass, and this assignment asks you about how you would approach several of those issues. You do not need to do any computing or analysis just yet, although the original data are available on the course web page by tomorrow if you would like to look at it. What you need to turn in for this assignment are written descriptions of how you determined what approach we should take toward analysis in this problem (Question 1) and what procedure or procedures you propose using to address the issues described (Questions 2 and 3). Be prepared to present your ideas to the class. You don't need to make slides or graphics or anything like that, just an oral description of what you propose. We may have a Part 2 of this Assignment where you will be asked to implement one or more of the ideas that arise from this first part of Assignment 3.

1. Discuss whether you believe it would be better to view this problem as one involving a Bayesian analysis of a mixture model, or as one we should approach with a model having several levels of prior distributions. *Hint: Read Assignment 2 – Background carefully before developing your answer to this question.*

2. Based on your answer to question 1, what would you include in a summary of your inferences associated with the problem. Be specific about types of summary information (e.g., five number summary or table of quantiles) and/or graphs and plots (e.g., scatterplot of y versus x, plot of empirical distribution of z). Indicate how those inferences can be used to address the goals of the Iowa DNR.

3. A statistical issue in the use of the hierarchical model developed in Assignment

2 – Background is the fixed value of the power $\theta = 1.0$ used in the analysis. Describe how you would conduct an an assessment of this modeling choice. In particular, there are two immediate alternatives to our choice. One is that a single value of θ should be adequate to reflect data behavior, but it's value should be something other than $\theta = 1.0$. In this case, we might assign θ a prior (similarly to α), but this question is not about what we might choose for that prior, only how we might assess the output of analysis of the hierarchical model (in files `LMBMCMC1.txt` and `LMBMCMC2.txt`) and the actual data (in file `LMBdat_for601.txt`) to determine whether we are motivated to include a prior for θ. The other possibility is that we might allow this power to have different values in different lakes, and make use of parameters $\{\theta_i : i = 1, \ldots, n\}$. These quantities would then need to be assigned a distribution, the parameters of which will be assigned a prior but, again, don't worry about what any of those distributions might be, only how we could determine whether there is evidence that a single value of θ, be that $\theta = 1$ or some other value, appears adequate or inadequate to reflect the behavior of the actual data.