



MT2002 Statistical Modeling

Assignment No: 04	CLO:
	Semester: Fall 2023
Due date: As of GCR date	Marks: 50

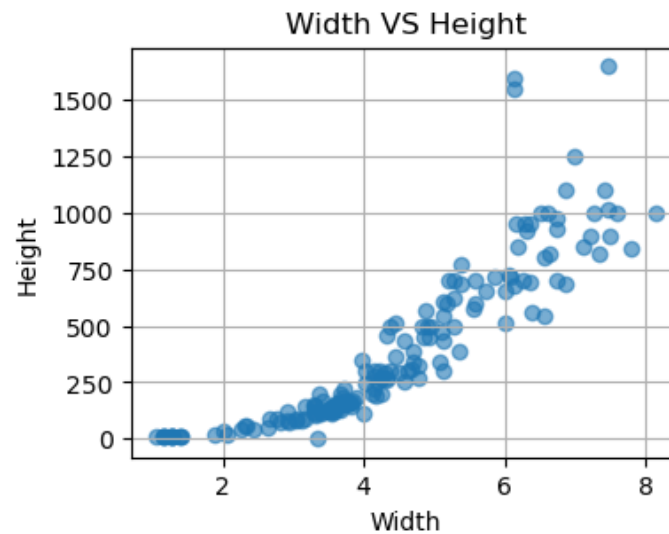
Instructions

1. **Plagiarized work will result in zero marks.**
2. **No retake or late submission will be accepted.**
3. The submission should be a SINGLE UNZIPPED NOTEBOOK submitted on google classroom.
4. This notebook should properly document what you did? How you did it? And the source-code for each part as well as the generated outputs.
5. Your submission file should be according to the following **format: id_section_A2** e.g., i22123456_A_A2. (Note: A2 in the end denotes Assignment 2).

Marks distribution for this assignment.

Models	10 marks each
Comparison	
First question	40 marks
Second question	

Examine the 'fish.csv' dataset from Kaggle, with Width as the independent variable and Height as the dependent variable. Here is the relationship between the two variables.



x = **Width**
y = **Height**

$\alpha \sim \text{Normal}(\mu_\alpha, \epsilon_\alpha)$
 $\beta \sim \text{Normal}(\mu_\beta, \epsilon_\beta)$
 $\epsilon \sim \text{Half-Normal}(\mu_\epsilon, \epsilon_\epsilon)$
 $\mu \sim \alpha + \beta * x$
 Likelihood $\sim \text{Normal}(\mu, \epsilon, \text{observed}=y)$

Linear Model

$\alpha \sim \text{Normal}(\mu_\alpha, \epsilon_\alpha)$
 $\beta_1 \sim \text{Normal}(\mu_{\beta_1}, \epsilon_{\beta_1})$
 $\beta_2 \sim \text{Normal}(\mu_{\beta_2}, \epsilon_{\beta_2})$
 $\epsilon \sim \text{Half-Normal}$
 $\mu \sim \alpha + \beta_1 * x + \beta_2 * x^2$
 Likelihood $\sim \text{Normal}(\mu, \epsilon, \text{observed}=y)$

Polynomial

Using ArviZ to compare the specified models mentioned above using information criteria such as LOO or WAIC as given in sample table below and next answer the questions.

Model	elpd_loo/WAIC	p_loo/WAIC	elpd_diff
Linear			
Polynomial ²			

Questions:

- Based on the above table, which model do you think is better from predictive accuracy point of view and why?
- Based on the above table, which model is complex and why?