Practice Class Test2 Template

2416963E

Question 1

Large samples of iron ore were mined from quarries "x" and "y", and each of the two samples were broken down into 10 smaller sub-samples for analysis. Each of these 20 sub-samples were sent to a chemical laboratory, and the percentage of iron in each sub-sample was measured. These data are stored in `practice1.csv`. Using bootstrapping, is there evidence in these data that the population mean iron percentage in each quarry is 35%, and are the population mean percentages different between the two quarries?

Question 2

(a) Using the data contained in `practice2.csv` build a regression model that adequately describes the `response` in terms of the potential explantory variables `X1`, `X2`, `X3` and `X4`. Your chosen model will therefore be the one that you believe best represents the `response`. Use only the theoretical confidence intervals generated under standard assumptions (which you should check) to identify the correct model. Note you **do not** need to consider interaction terms.

(b) Construct a table of **all** the possible linear models (without interactions or transformations) that could be fitted to the `response` variable in `practice2.csv`. In the table include the R^2_{adj} and AIC values for model comparisons. Do these measures lead you to the same conclusion about the model that best represents the data as in part (a)? Note: you are **not** required to check assumptions for each of these models in this task.