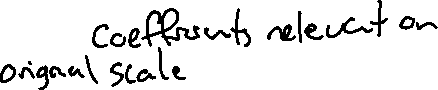
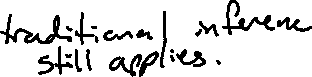
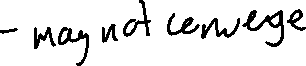
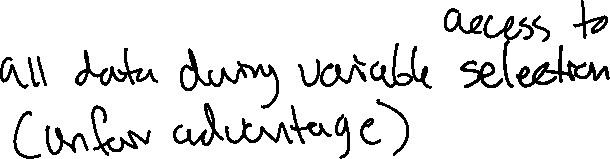
(4.2) Suppose you have a linear model which demonstrates heteroskedasticity in the residuals. Suppose that the transformation of the response variable effectively removes this heteroskedasticity. Why might you prefer this transformation to a weighted least squares approach? Why might a weighted least squares approach be preferred?



(4.2) Name one risk associated with determining parameters in non-linear models?



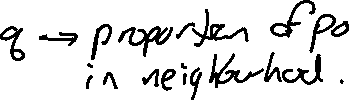
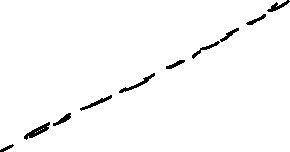
(4.3) When variable selection is applied before cross validation, will the resulting cross validated errors be an over-estimate or under-estimate of the true error we would see on new data?



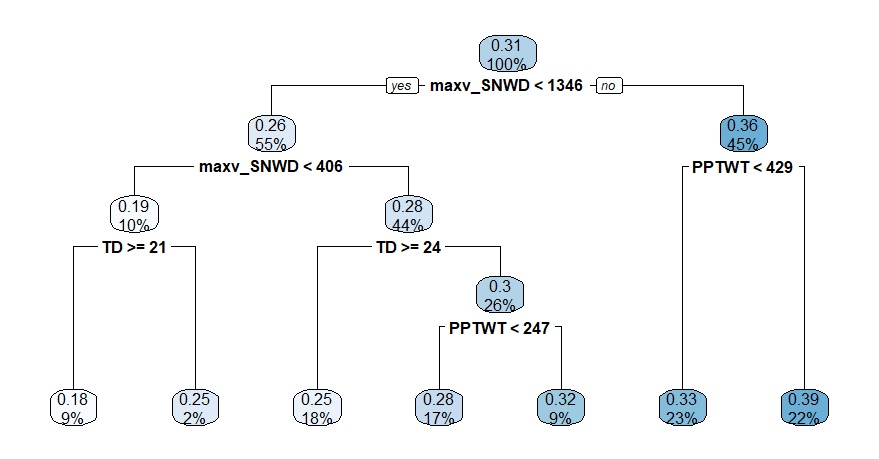
(4.4) Two loess curves were fit to the same data, each using different smoothing parameters (q). Based on the visualization below, which curve (red or blue) has a larger smoothing parameter?

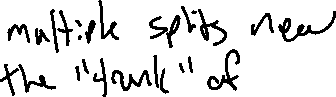
A close up of a map

Description automatically generated

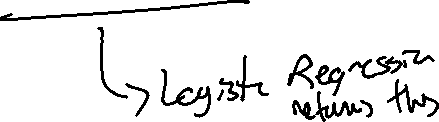
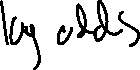
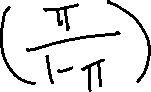


(4.4) Given the provided tree, which variable is most important in predicting snow density?

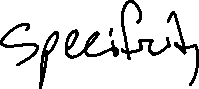
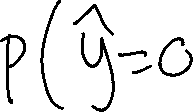
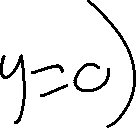
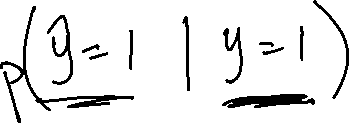
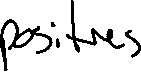




(5.1) What is the difference between a log-odds and a probability?



(5.1) What is the appropriate definition of “sensitivity” in the context of Logistic regression?



(6.1) What does this plot provide evidence for? Why is that a problem?

