Statistics 745

Assignment 2

Name:_____

1. Let y be the response and X be the data such that:

$$X = \begin{pmatrix} 1 & 0 \\ 1 & 0 \\ 1 & 0 \\ 1 & 0 \\ 0 & 1 \\ 0 & 1 \\ 0 & 1 \\ 1-\tau & \tau \end{pmatrix} \quad y = \begin{pmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \\ y_5 \\ y_6 \\ y_7 \\ y_8 \end{pmatrix},$$

   with $\tau \in \{0, 1\}$.

   (a) Determine $\hat{\beta}^{(ls)}$ without the intercept as a function of τ.

   (b) Determine the k-NN classifier with $k = 2$ as a function of τ.

   (c) Compare k-NN to $X\hat{\beta}^{(ls)}$. Explain.

2. **Individual Effort**:

   (a) Implement your own least squares classifier in R.

      i. Apply it to the $sin(x)$ regression example provided as supplementary material.

      ii. Apply it to the simulated classification example.

   (b) Implement your own k-NN function.

      i. Apply it to the $sin(x)$ regression example provided as supplementary material. Investigate the effect of k.

      ii. Apply it to the simulated classification example. Investigate the effect of k.