# MATH 282B – Homework 6

**Problem 1.**

   A. Write a function $\mathsf{piecewiseConstant}(\mathsf{x},\ \mathsf{y},\ \mathsf{q}\ =\ 10,\ \mathsf{plot}\ =\ \mathsf{TRUE})$ taking in a one dimensional predictor variable $x$ with values in $[0, 1]$ and a response $y$, and fits a piecewise constant model (by least squares) on $q$ intervals of equal length partitioning the unit interval, with the option of producing a scatterplot with the fit overlaid.

   B. Apply your function to some dataset of your choice, either synthetic or real. (If you choose a real dataset, then provide the source.) Produce a single scatterplot, with lines corresponding to the fit with $q = 5$ (blue), $q = 10$ (green), and $q = 20$ (red). Add a legend.

**Problem 2.** Consider a binary response (values in $\{0, 1\}$ without loss of generality) and a two-dimensional predictor where

$$\mathbb{P}(y = 1 \mid \mathbf{x}) = \exp(-\|\mathbf{x}\|) \qquad \mathbf{x} \sim \mathcal{N}(0, \mathbf{I})$$

We work with a sample of size $n$ from this model, where $n \in \{50, 100, 200, 500, 1000\}$.

   A. Fit a simple logistic model. Explain why the performance is bound to be mediocre.

   B. Fit a higher order polynomial logistic model. Consider several variants: degree 2 without interactions, degree 2 with interactions, and higher degree models. Track the average performance by averaging over $R = 20$ repeats. Produce a plot or two summarizing the results.