**Problem 1**

Question 1

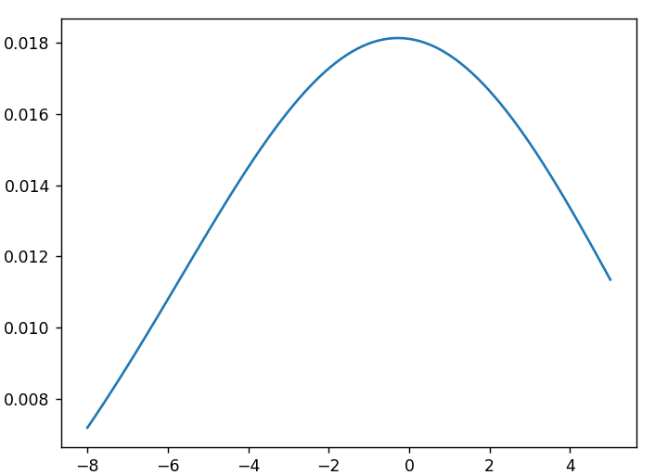
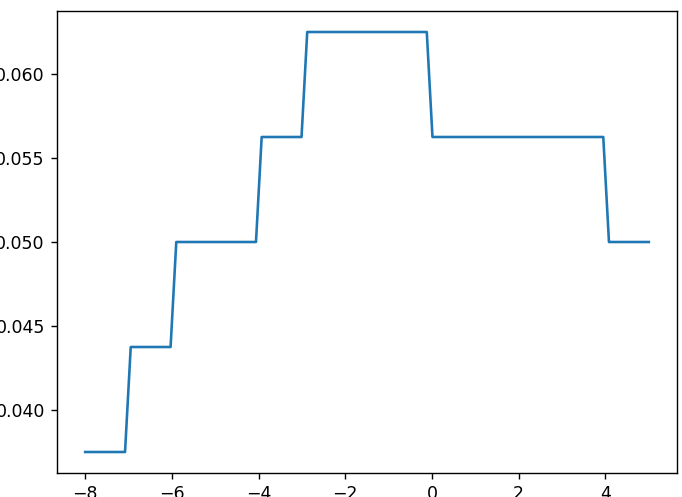
Notice that the risk is:

Substituting 0-1 loss function we get:

We want to minimize the risk, which means to maximize the probability . Notice that:

Notice that when , , and therefore we're simply counting the number of votes of each class.

Question 2

Here's a drawing of the probability estimation:

We assume the distribution is gaussian. Here are the MLE estimators: