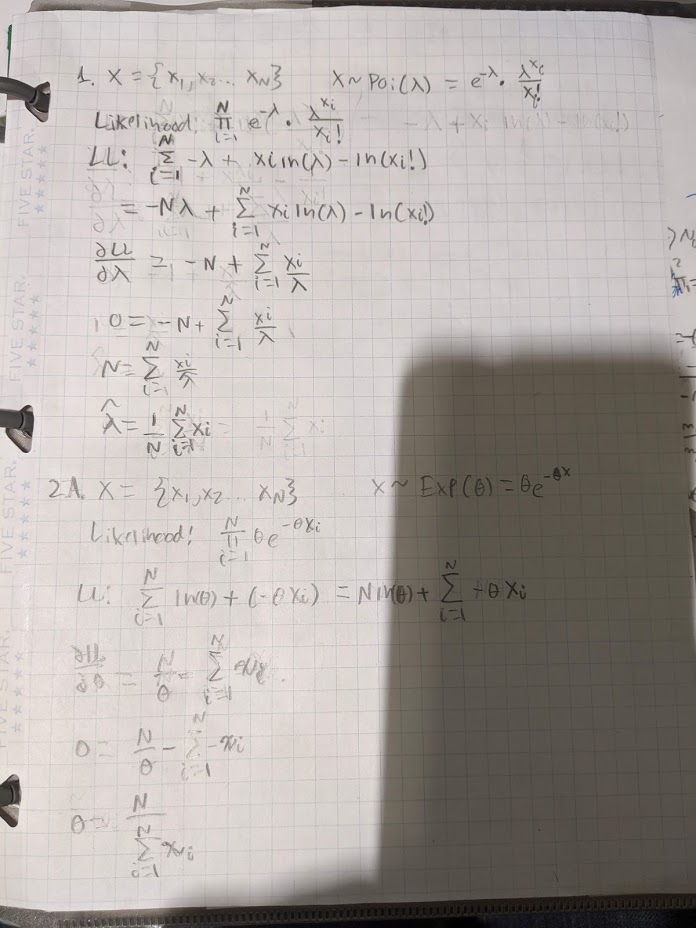
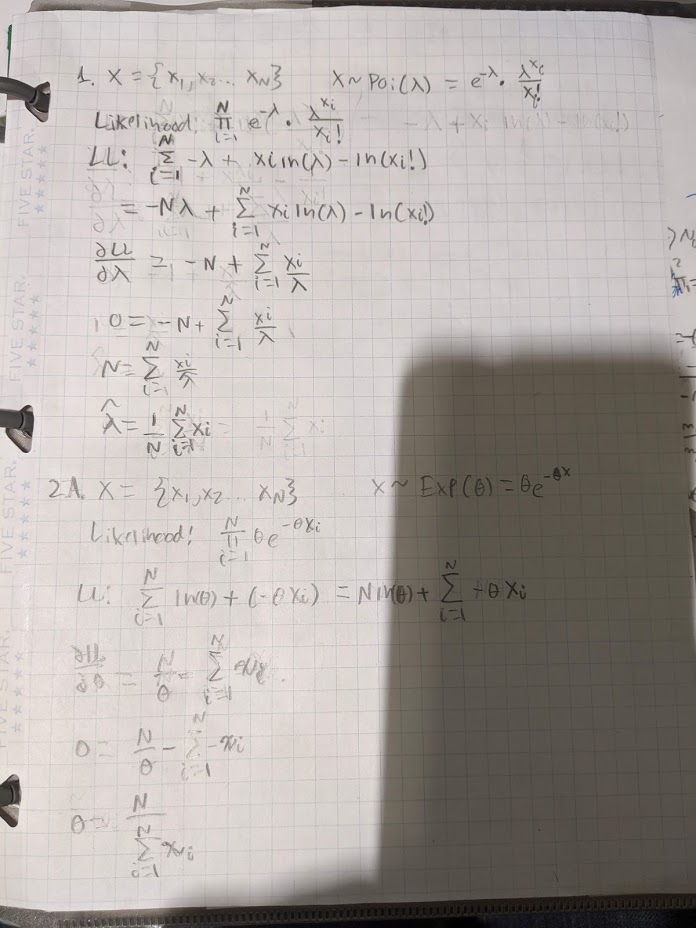
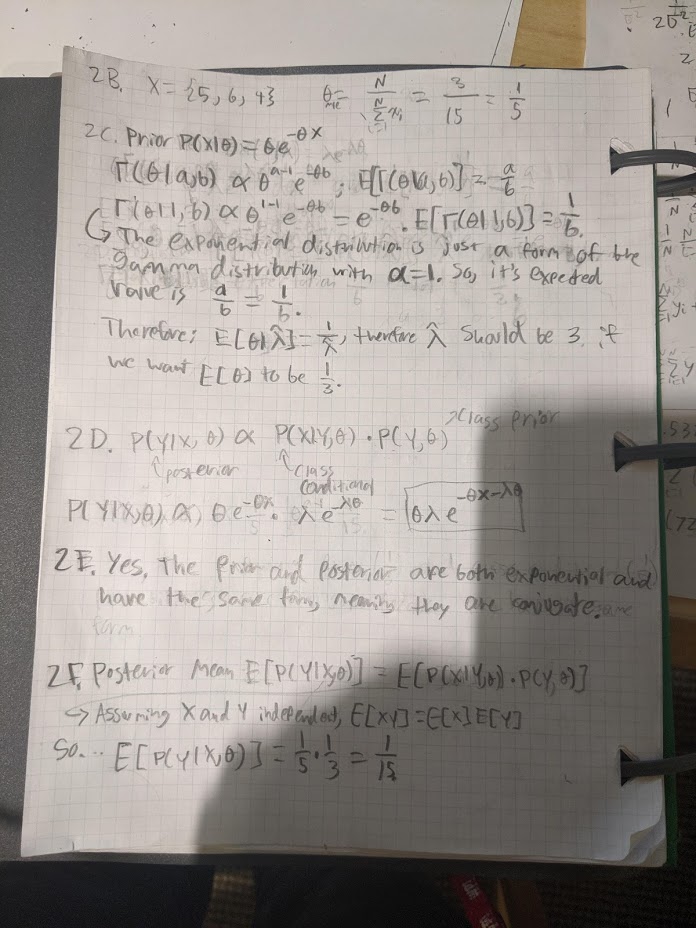
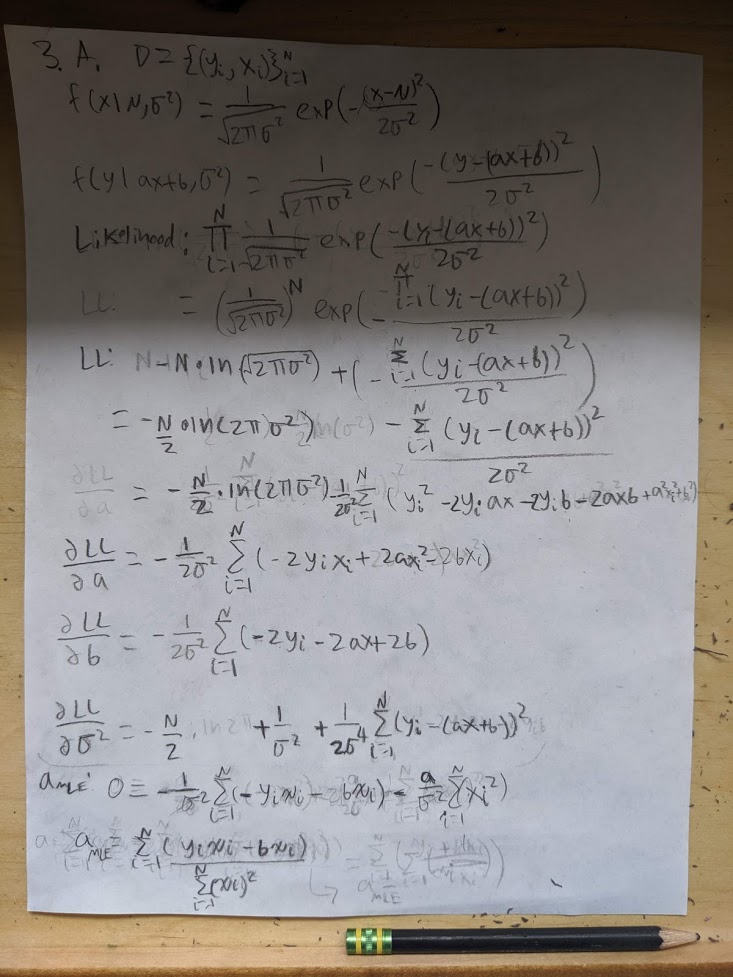
**Problem 1**

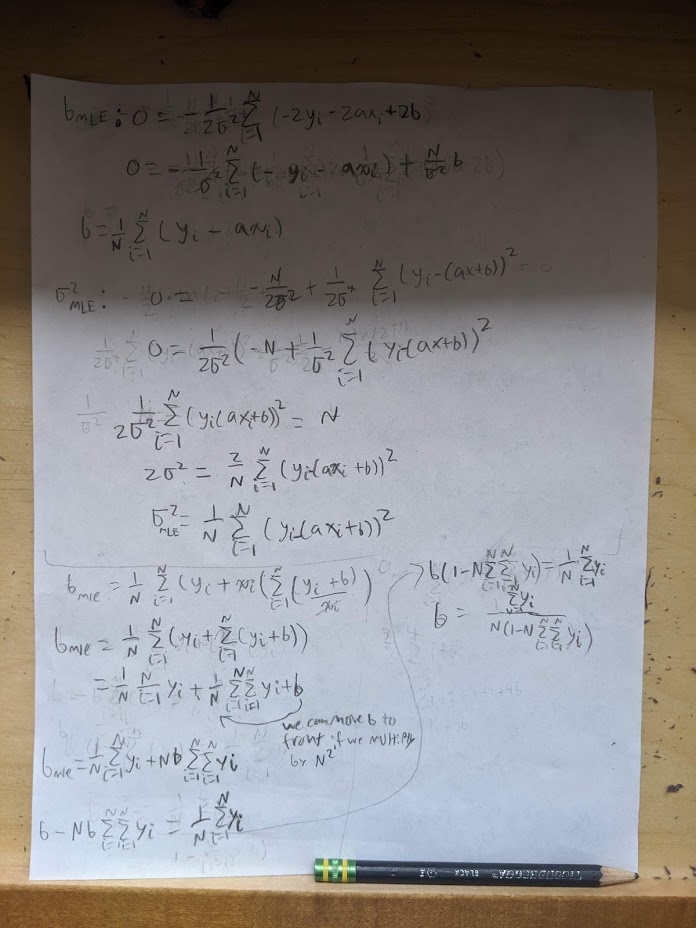


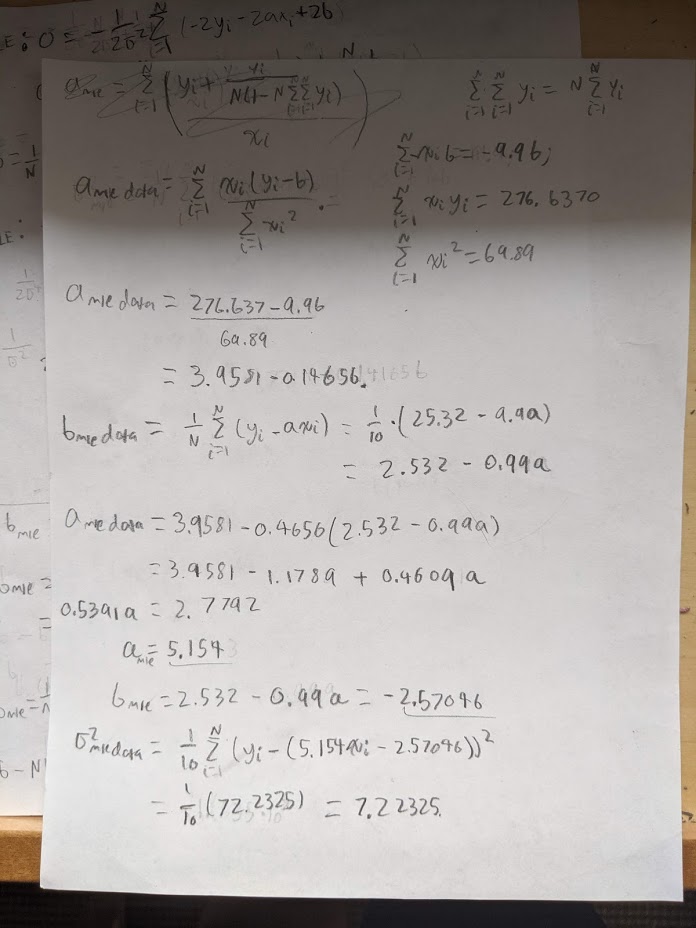
**Problem 2**



**Problem 2 Continued**

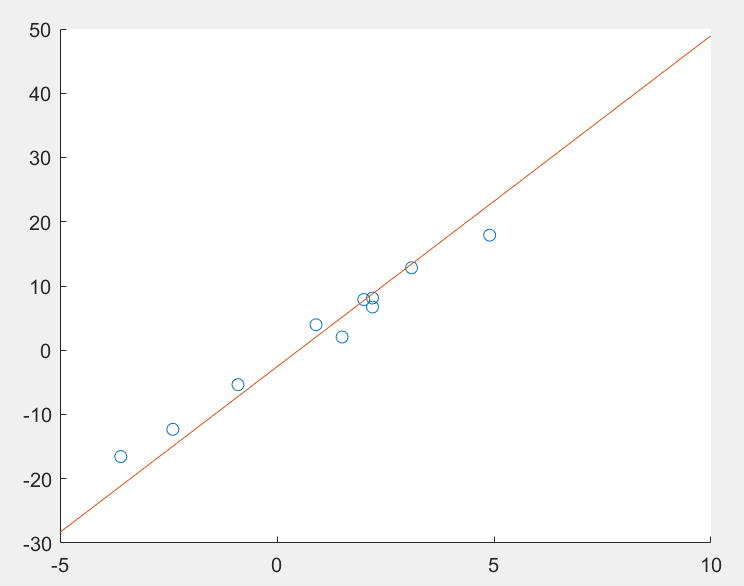
**Problem 3**

**PROBLEM 3A CONTINUED**

**PROBLEM 3B**

Final values: a = 5.154, b = -2.57046, 𝜎2 = 7.22325.

**PROBLEM 3B CONTINUED**

****

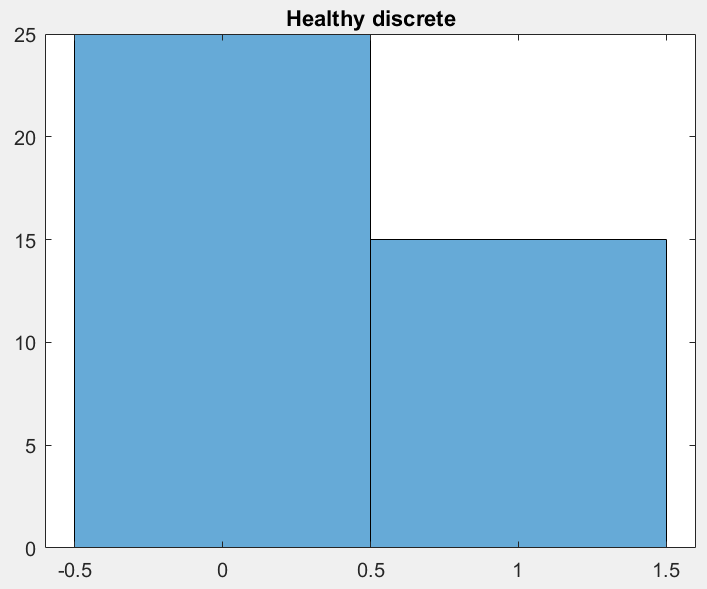
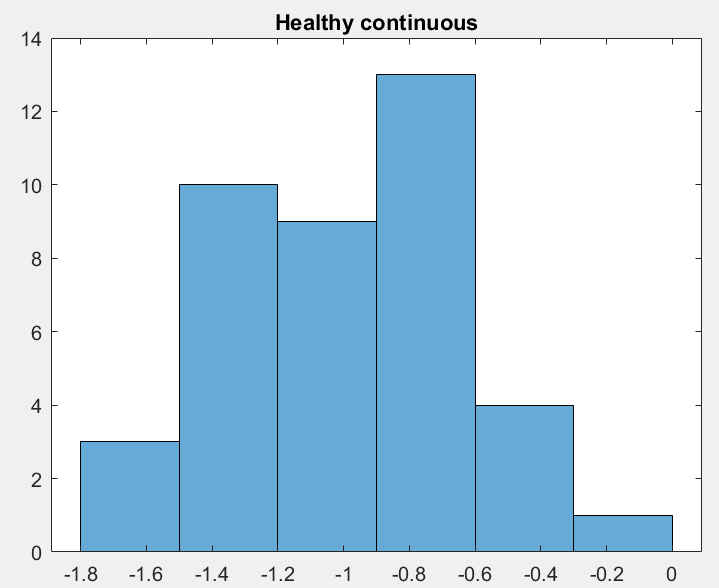
y = 5.154x - 2.57046

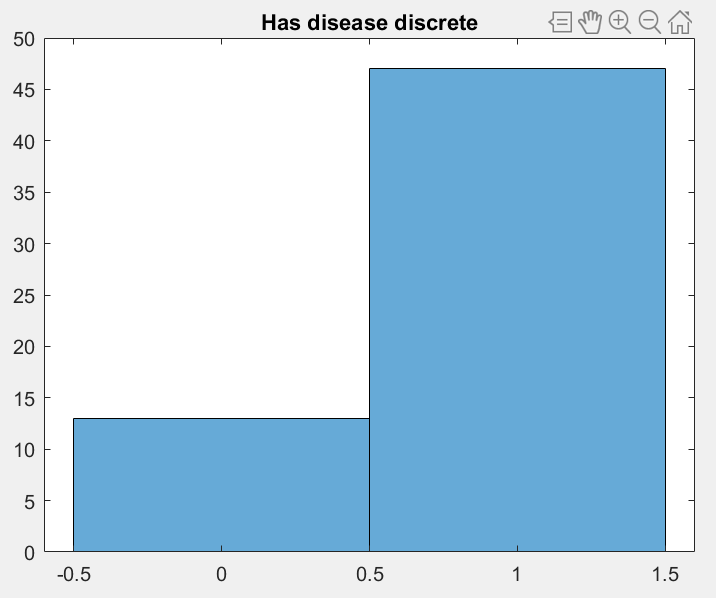
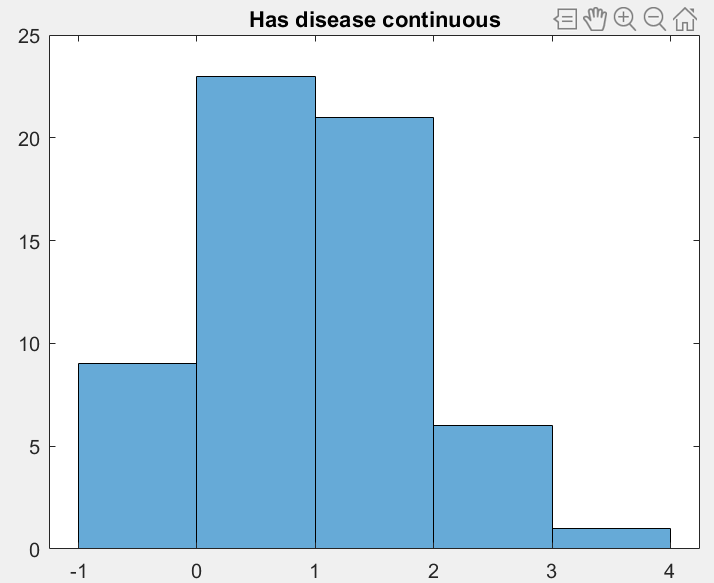
**PROBLEM 3C**. I suggest normal priors for both A and B to make them conjugate with the prior for 𝜎2.

**(3D posted to Canvas, not required)=**

**Problem 4**

|  |  |  |  |
| --- | --- | --- | --- |
|  |  | Has disease | Healthy |
|  |  | 1 | 0 |
| Has indicator | 1 | 47 | 15 |
| No indicator | 0 | 13 | 25 |

1. 



1. Basic classifier for discrete

|  |  |  |
| --- | --- | --- |
|  | Diseased | Healthy |
| Classified Correctly | 47 (TP 78.3%) | 25 (TN 62.5%) |
| Misclassified | 13 (FP 21.7%) | 15 (FN 37.5%) |

**PRIORS AND CONDITIONAL DISTRIBUTION PARAMETERS:**

P(y='disease'|x='1'): 0.47

P(y='healthy'|x='1'): 0.15

P(y='disease'|x='0'): 0.13

P(y='healthy'|x='0'): 0.25

P(y='disease'): 0.6

P(y='healthy'): 0.4

(probability of having disease ~ Bernoulli with p = 0.6)

1. Basic classifier for continuous

|  |  |  |
| --- | --- | --- |
|  | Diseased | Healthy |
| Classified Correctly | 55 (TP 91.7%) | 39 (TN 97.5%) |
| Misclassified | 5 (FP 8.3%) | 1 (FN 2.5%) |

Distribution for healthy: ~ Normal with mu = -1, sigma =0.367

Distribution for diseased: ~ Normal with mu = 0.941, sigma = 0.887

1. Naïve Bayesian Classifier with both:

|  |  |  |
| --- | --- | --- |
|  | Diseased | Healthy |
| Classified Correctly | 56 (TP 93.3%) | 38 (TN 95%) |
| Misclassified | 4 (FP 6.7%) | 2 (FN 5%) |

TP:

**Continuous variable:**

Distribution for healthy: Normal with mu = -1, sigma =0.367

Distribution for diseased: Normal with mu = 0.941, sigma = 0.887

Discrete variable:

1. The basic classifier using the discrete variable learned that data points with the indicator were more likely to have the disease than not, and therefore predicted that all data with the indicator had the disease. This led to false positives and false negatives as some data points had the indicator but not the disease and vice versa. The discrete basic classifier misclassified 28% of the data. The basic classifier using the continuous variable did better, because the classifier’s normal distribution for diseased patients reflected the fact that nearly all patients who had the disease also had a positive continuous indicator value, which greatly reduced the number of false negatives. However, there were some false positives, because some patients with negative indicator values were also diseased. The continuous classifier misclassified only 6% of the data. The Naïve Bayesian Classifier, which took into account the probabilities of both indicators, performed the same overall as the basic continuous classifier, and also misclassified 6% of the data. However, the Naïve Bayesian had fewer false positives and more false negatives. In this particular problem, the basic continuous classifier is probably the best option because its errors are mostly false positives, which in the case of detecting a disease, is probably better than having false negatives.

**Problem 5**

X1:

Probability class 1: -1.0991

Probability class 2: -2.2599

Probability class 3: -1.2599

X2:

Probability class 1: -1.0991

Probability class 2: -1.0932

Probability class 3: -2.7599

**Conclusions:**

1. **X1 belongs in class 1.**
2. **X2 belongs in class 2.**

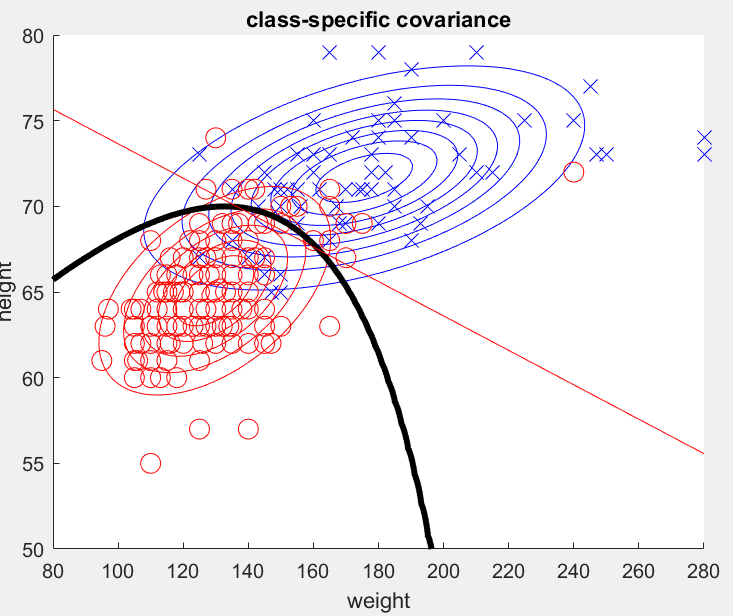
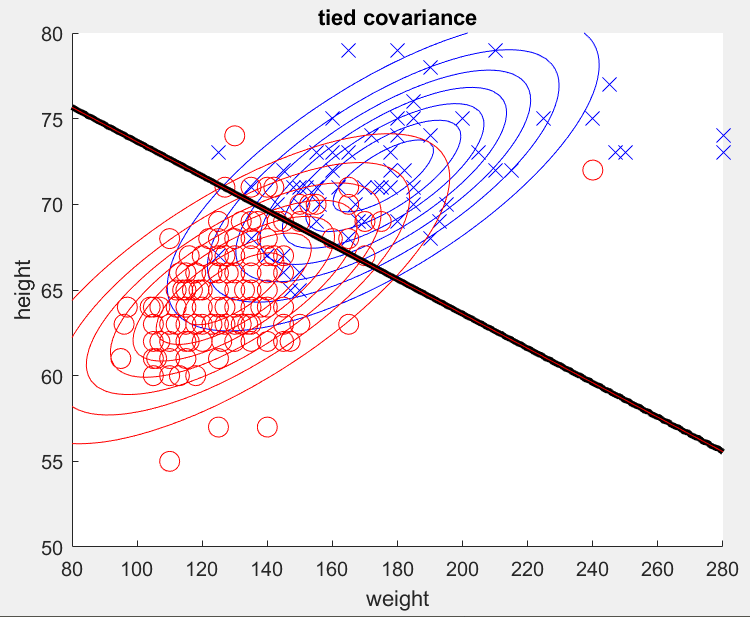
**Problem 6**

QDA Statistics:

|  |  |  |  |
| --- | --- | --- | --- |
| Class | Classified Correctly | Classified Incorrectly | Misclassification Rate |
| Class 1 | 64 | 9 | 12.33% |
| Class 2 | 121 | 16 | 11.68% |

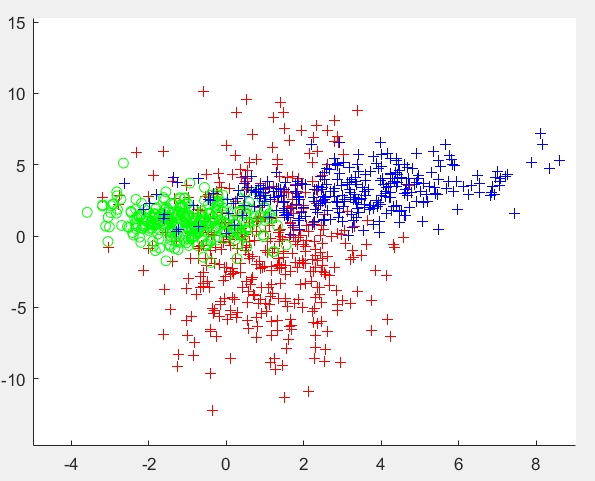
LDA Statistics:

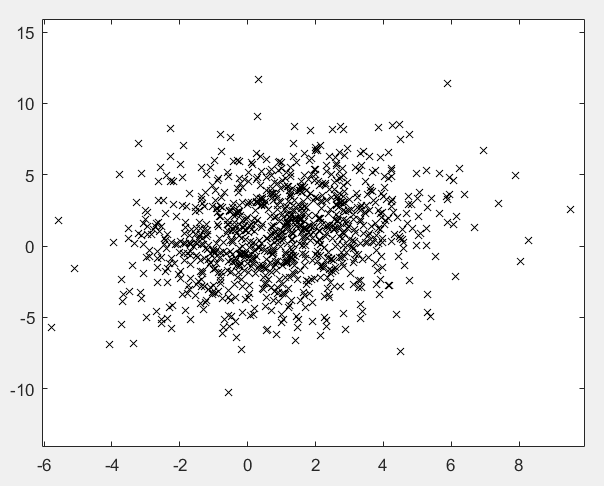
|  |  |  |  |
| --- | --- | --- | --- |
| Class | Classified Correctly | Classified Incorrectly | Misclassification Rate |
| Class 1 | 64 | 9 | 12.33% |
| Class 2 | 120 | 17 | 12.41% |

****

**Note: I flipped the x and y axes in order to obtain a valid polynomial function for the QDA curve. This shouldn’t affect the validity of my calculations, but the graphs here and in the code are rotated 90 degrees from the originals.**

**Problem 7**

****



The difference between these two distributions is that the first is composed of 3 fairly distinct shapes put together, whereas the second is more of a single blob or cloud shape. It’s like the difference between a wheel with 3 spokes and a solid wheel; they might both have the same center (mean) and diameter (covariance), but their masses are distributed differently.