Homework 5

Adam Karl

March 4, 2021

1 Problem 1

b.
training data confusion matrix:

0.4879	0.1651
0.1410	0.2059

training misclassification error = 0.3061
testing data confusion matrix:

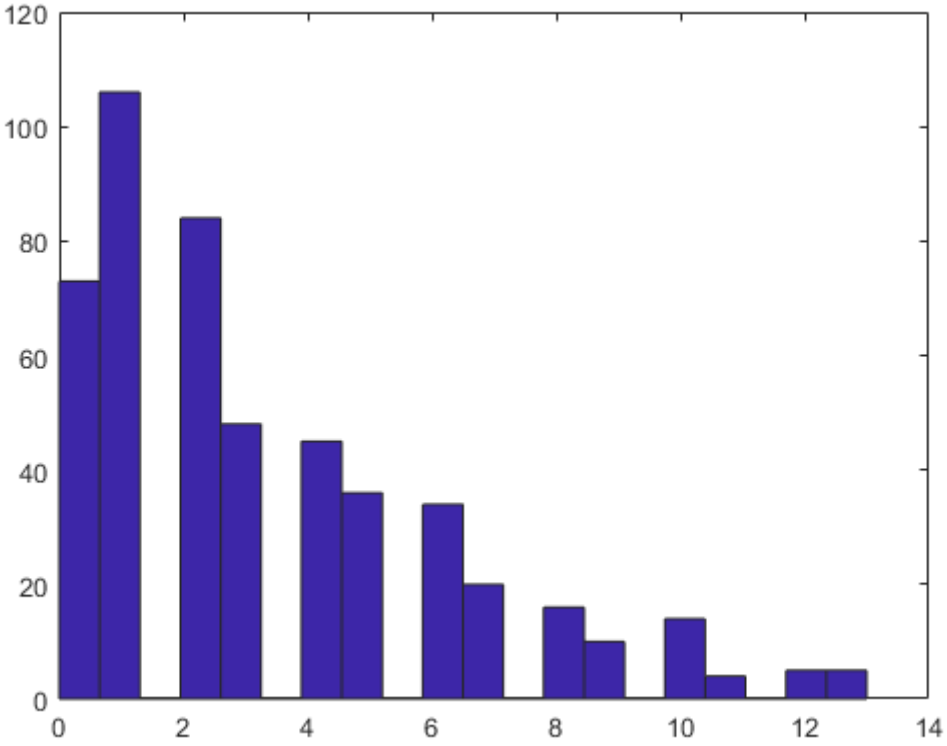
0.5284	0.0961
0.1747	.2009

testing misclassification error = 0.2707
testing sensitivity = 0.7516
testing specificity = 0.6765
c.

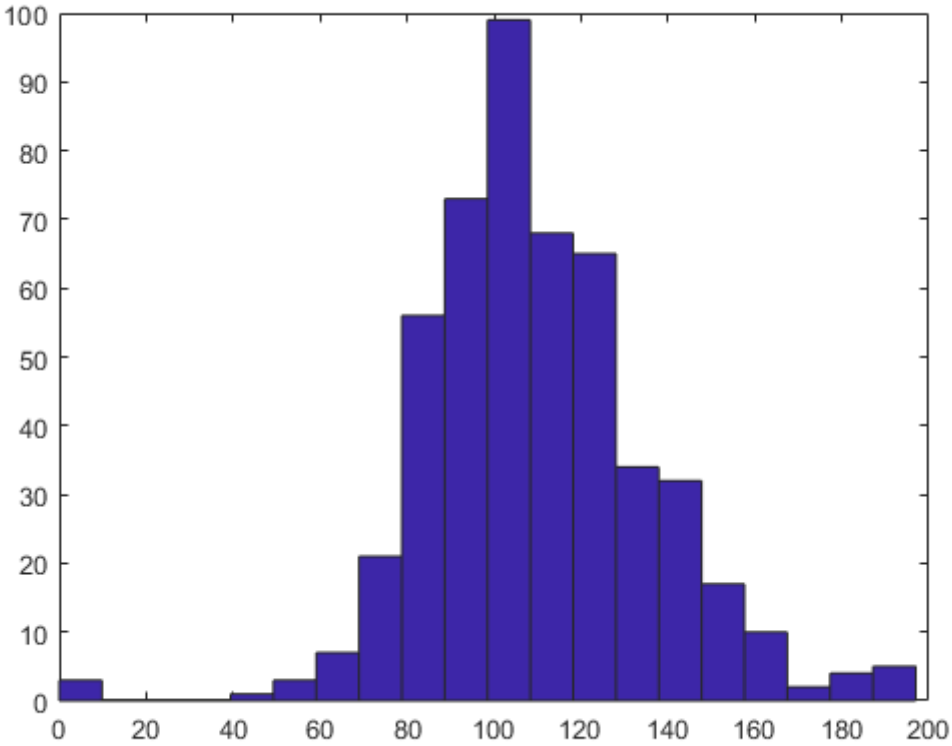
- 10,000 epochs
  - training misclassification error = 0.2839
  - testing misclassification error = 0.2533
- initial weights all 0.5
  - training misclassification error = 0.3061
  - testing misclassification error = 0.2707
- initial weights all 0
  - training misclassification error = 0.2542
  - testing misclassification error = 0.2358
- learning schedule = 1
  - training misclassification error = 0.2913
  - testing misclassification error = 0.3057
- learning schedule = .1
  - training misclassification error = 0.2820
  - testing misclassification error = 0.3057

With my experimentation the minimum error was achieved with initial weights set to 0 but all other settings default. I was able to get a training misclassification error of 0.2542 and a testing misclassification error of 0.2358.

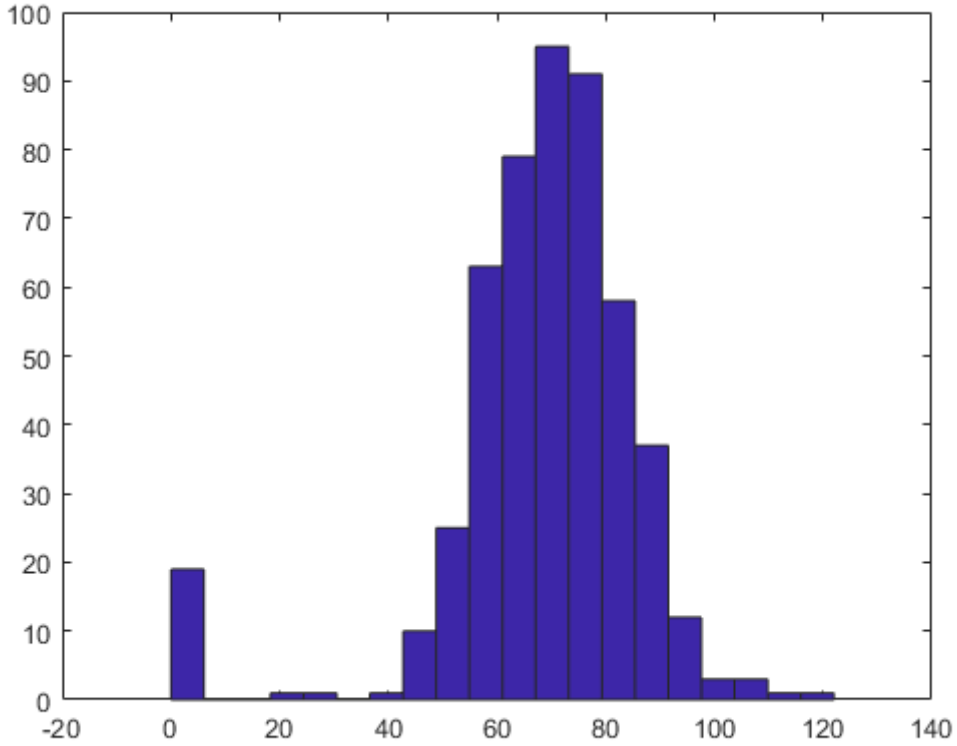
1.1 Problem 2.1



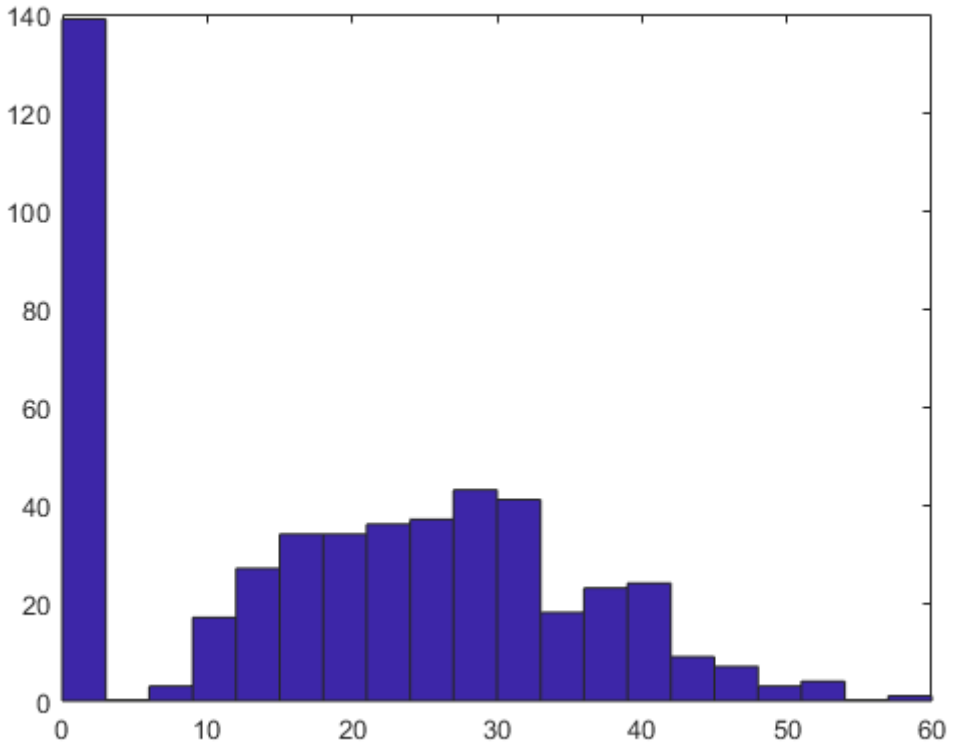
class 0, variable 1



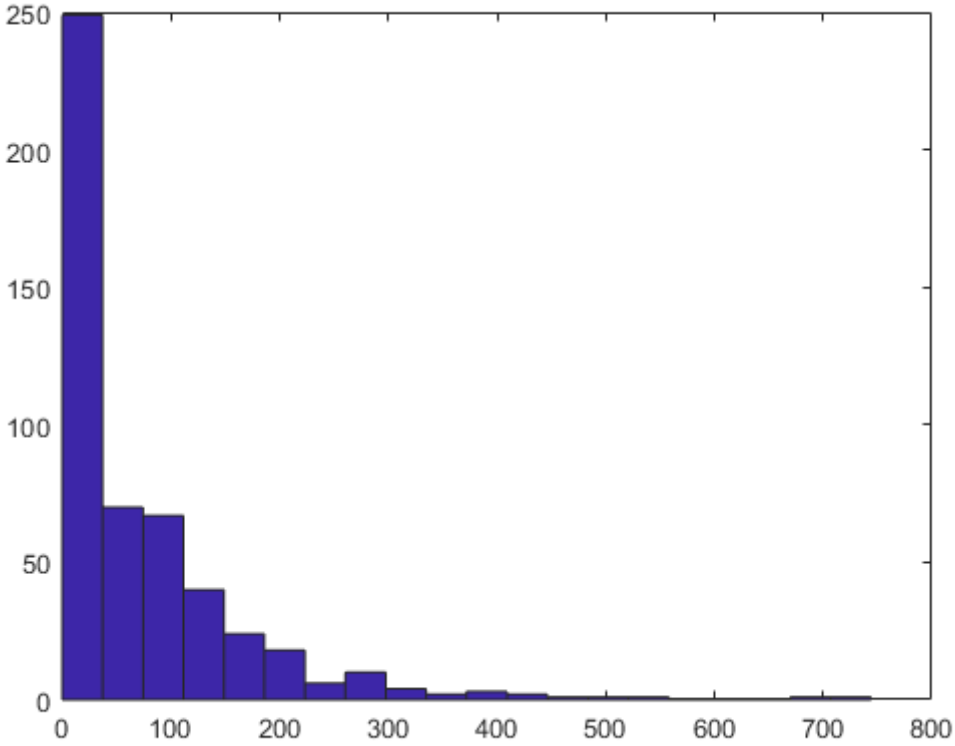
class 0, variable 2



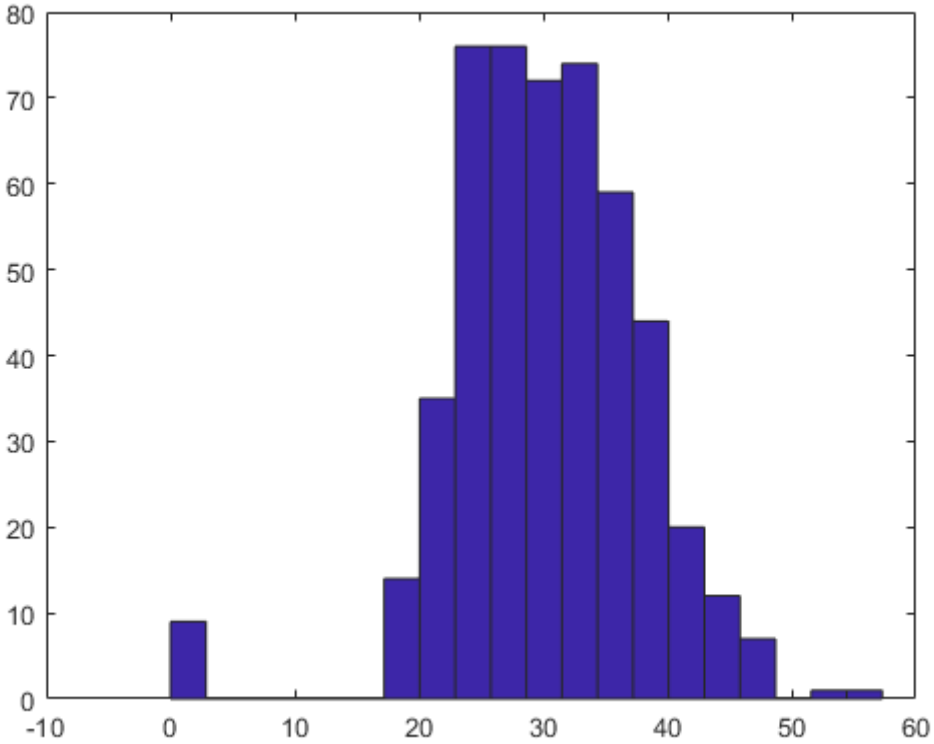
class 0, variable 3



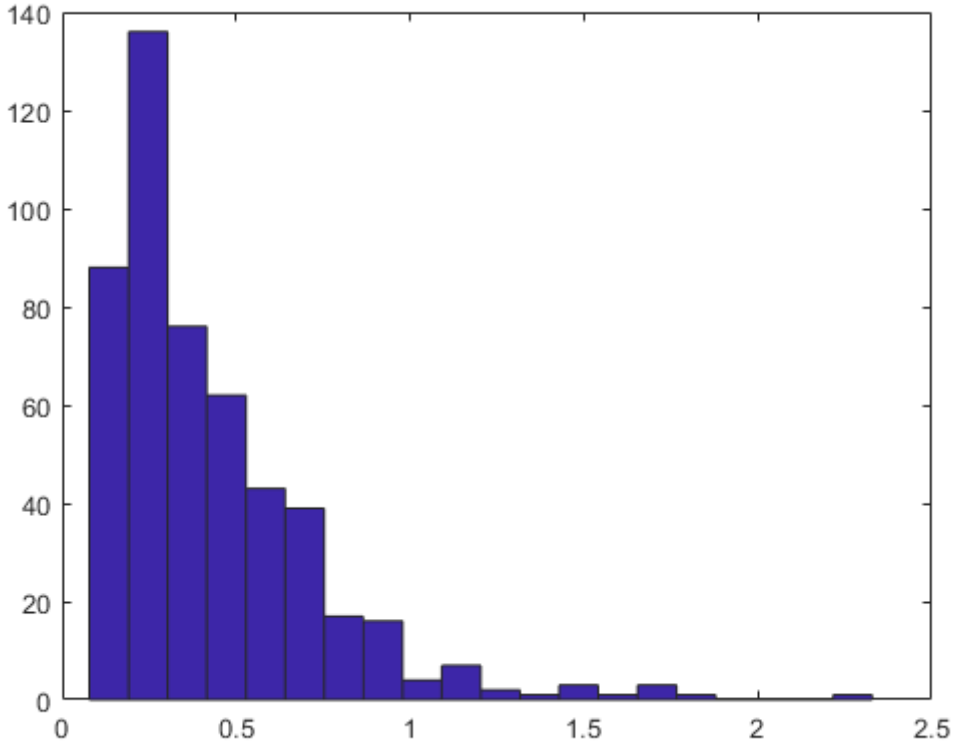
class 0, variable 4



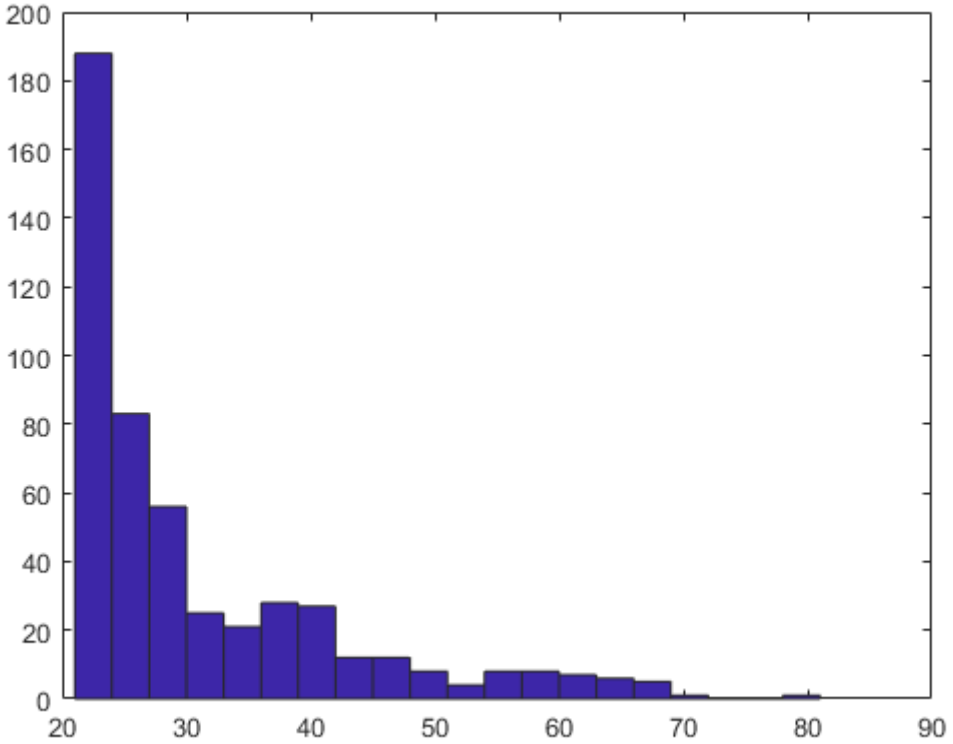
class 0, variable 5



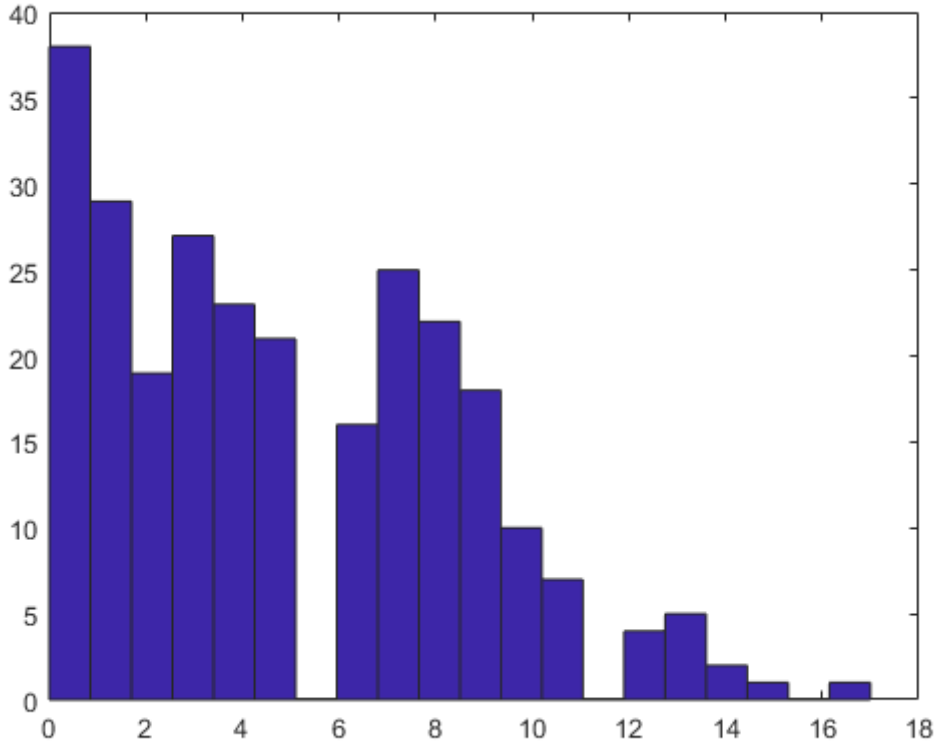
class 0, variable 6



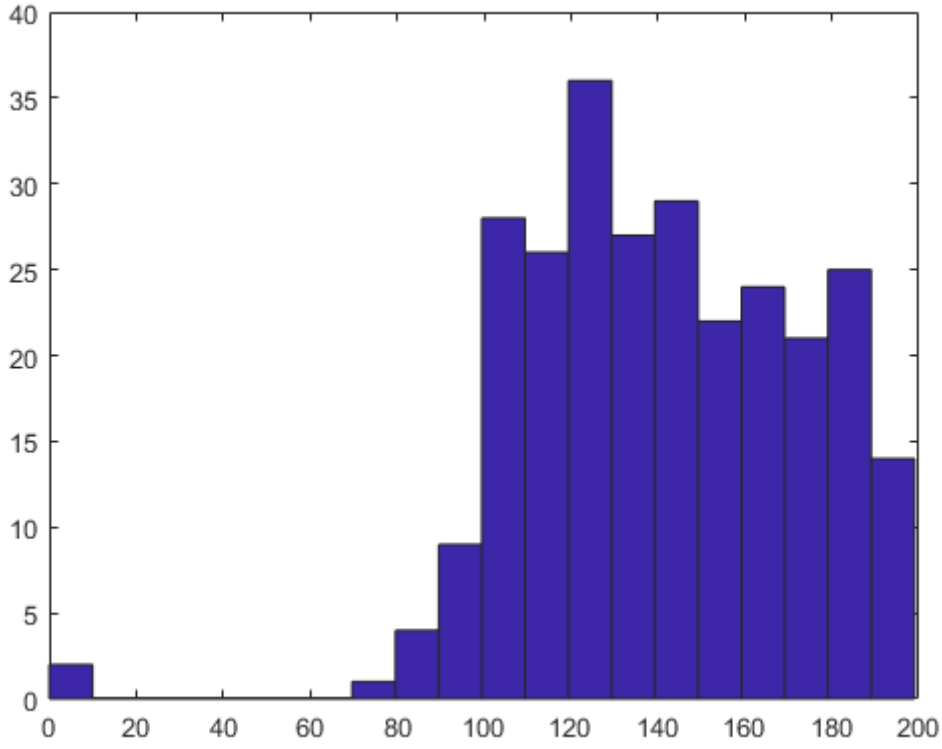
class 0, variable 7



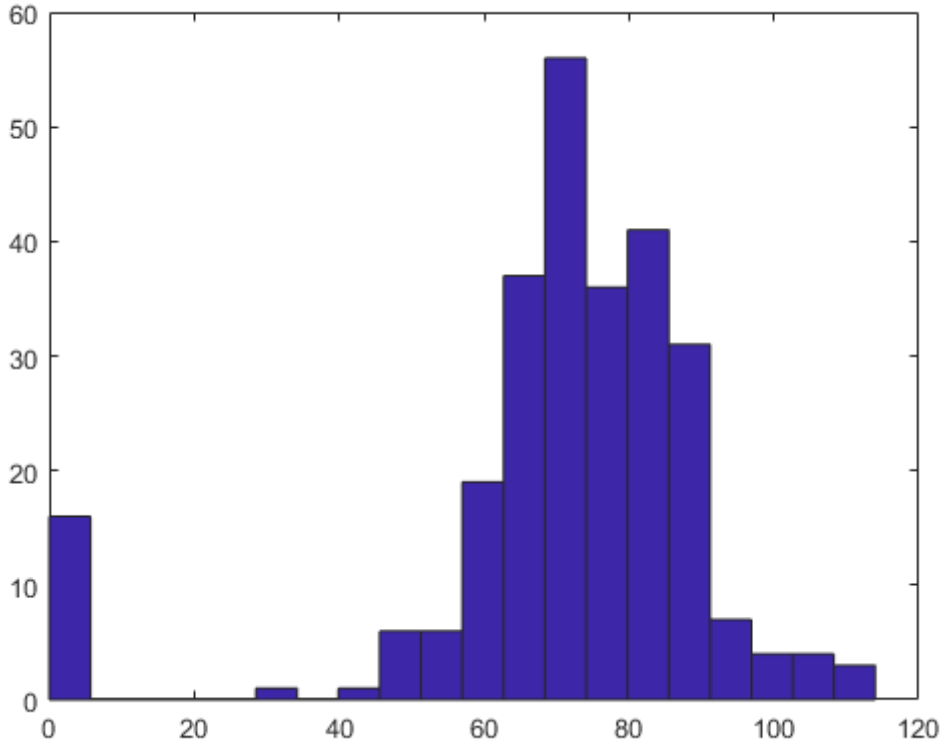
class 0, variable 8



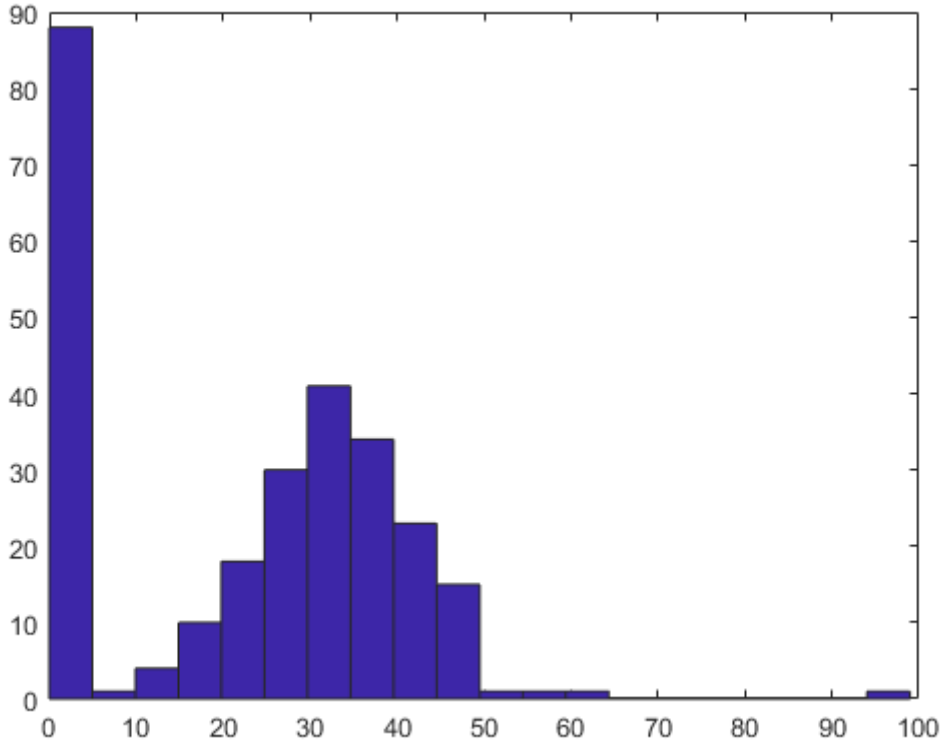
class 1, variable 1



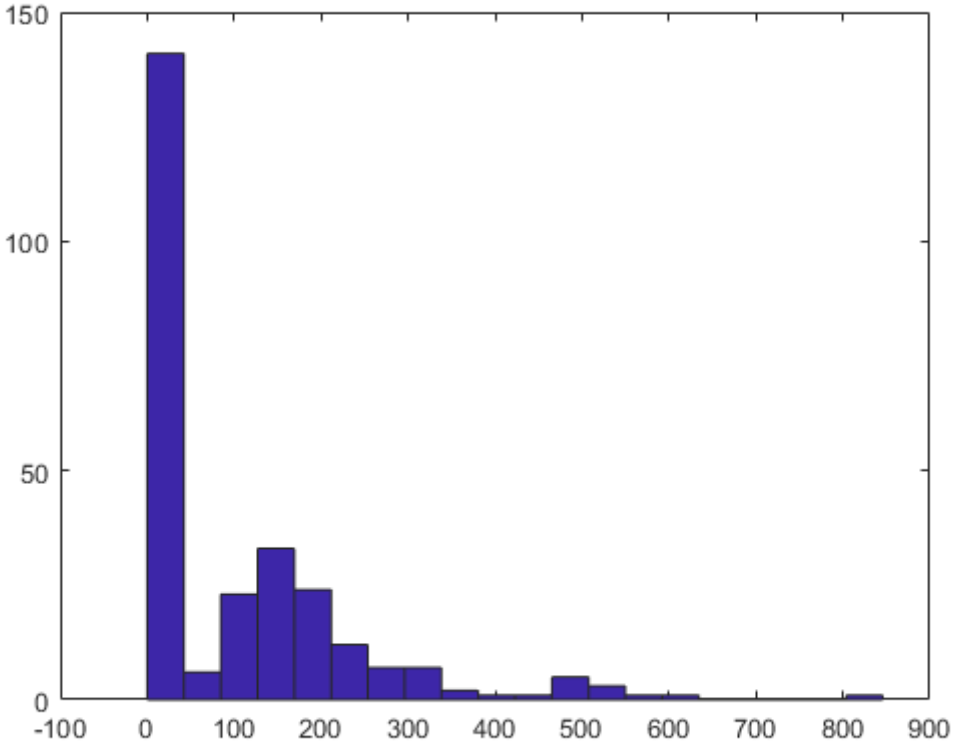
class 1, variable 2



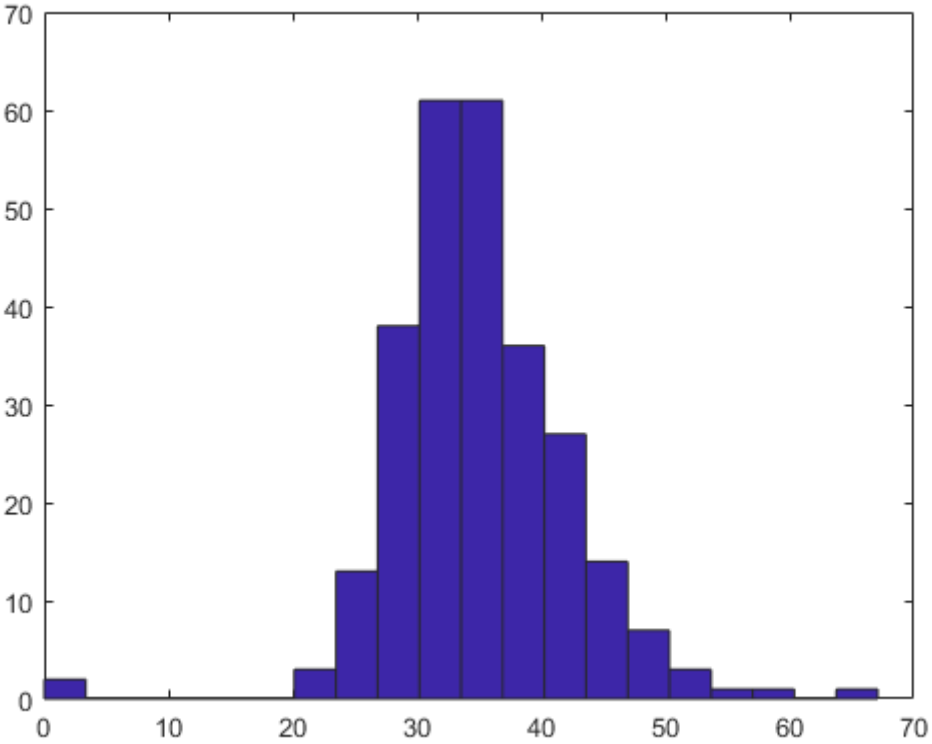
class 1, variable 3



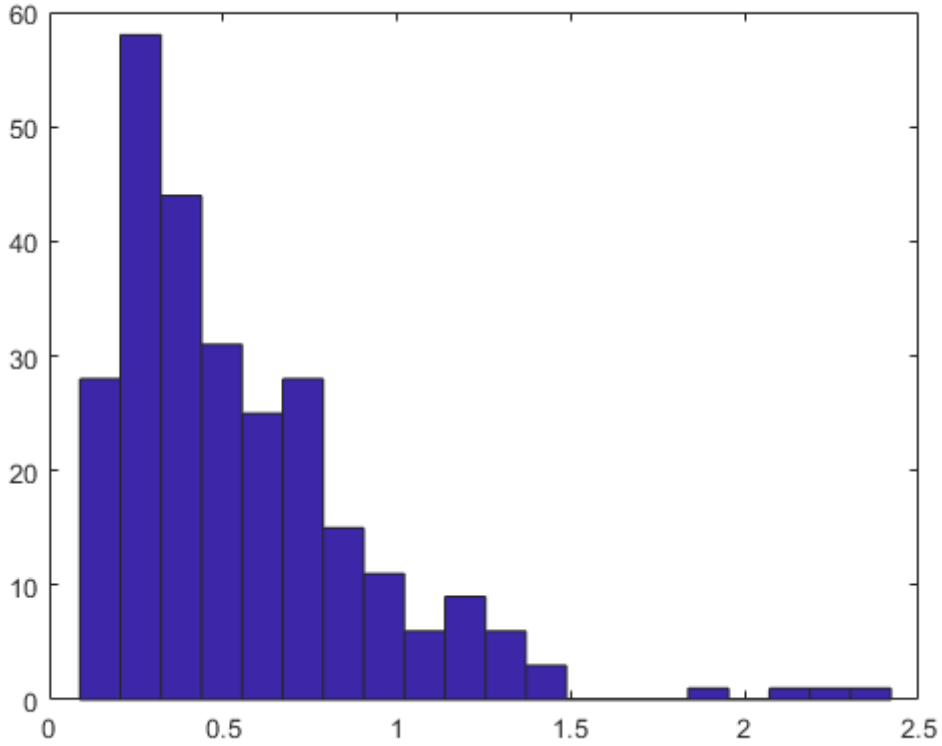
class 1, variable 4



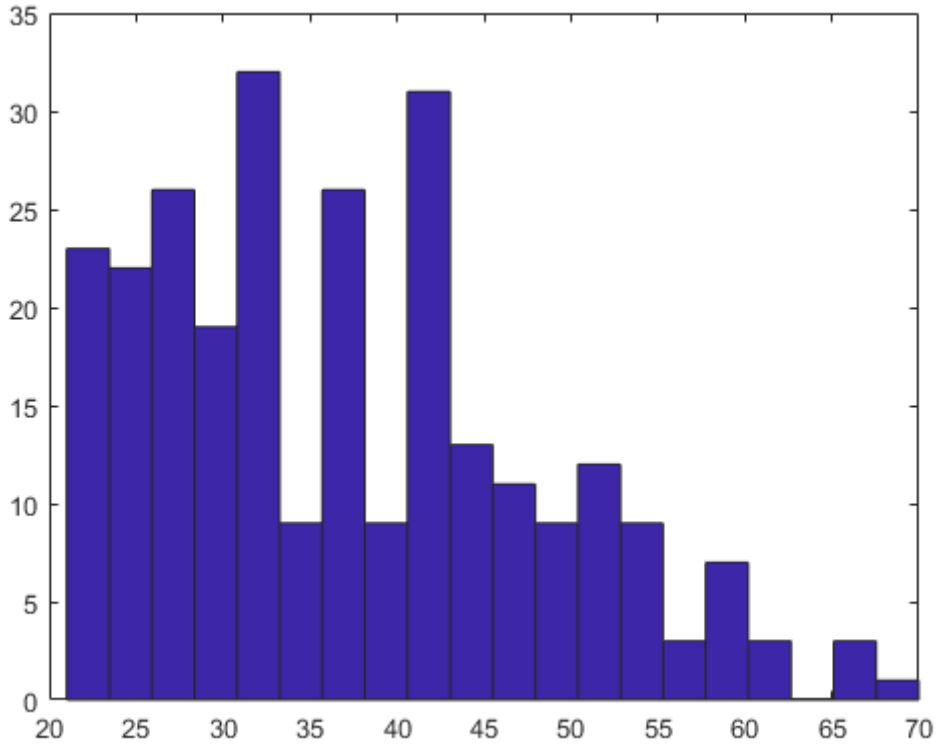
class 1, variable 5



class 1, variable 6



class 1, variable 7



class 1, variable 8

b.

- class 0

  - variable 1: binomial
  - variable 2: normal
  - variable 3: normal
  - variable 4: possibly loose normal, except for a lot of 0s which throw it off
  - variable 5: gamma
  - variable 6: normal
  - variable 7: gamma
  - variable 8: gamma
- class 1
  - variable 1: loose gamma
  - variable 2: very loose normal
  - variable 3: normal, except a fair amount of 0s which throw it off
  - variable 4: possibly normal, except for a lot of 0s which throw it off
  - variable 5: possibly normal, except for a lot of 0s which throw it off
  - variable 6: normal
  - variable 7: gamma
  - variable 8: very loose normal

1.2 Problem 2.2

- CLASS 0 EXPFIT
  - variable 1
    - muhat = 3.2419
  - variable 2
    - muhat = 109.6254
  - variable 3
    - muhat = 67.5339
  - variable 4
    - muhat = 19.7316
  - variable 5
    - muhat = 67.7168
  - variable 6
    - muhat = 30.3059
  - variable 7
    - muhat = 0.4164
  - variable 8
    - muhat = 31.1032
- CLASS 1 EXPFIT
  - variable 1
    - muhat = 4.7100
  - variable 2
    - muhat = 141.3950
  - variable 3
    - muhat = 70.1900
  - variable 4

        * muhat = 22.9350
    - variable 5
        * muhat = 103.7200
    - variable 6
        * muhat = 35.2580
    - variable 7
        * muhat = 0.5491
    - variable 8
        * muhat = 37.1200
- CLASS 0 NORMFIT
    - variable 1
        * mu = 3.2419
        * sigma = 2.9644
    - variable 2
        * mu = 109.6254
        * sigma = 26.2304
    - variable 3
        * mu = 67.5339
        * sigma = 18.6683
    - variable 4
        * mu = 19.7316
        * sigma = 14.5828
    - variable 5
        * mu = 67.7168
        * sigma = 91.6702
    - variable 6
        * mu = 30.3059
        * sigma = 7.7258
    - variable 7
        * mu = 0.4164
        * sigma = 0.2906
    - variable 8
        * mu = 31.1032
        * sigma = 11.3830
- CLASS 1 NORMFIT
    - variable 1
        * mu = 3.2419
        * sigma = 2.9644
    - variable 2
        * mu = 109.6254
        * sigma = 26.2304
    - variable 3
        * mu = 67.5339
        * sigma = 18.6683
    - variable 4

    - mu = 19.7316
    - sigma = 14.5828
  - variable 5
    - mu = 67.7168
    - sigma = 91.6702
  - variable 6
    - mu = 30.3059
    - sigma = 7.7258
  - variable 7
    - mu = 0.4164
    - sigma = 0.2906
  - variable 8
    - mu = 31.1032
    - sigma = 11.3830

1.3 Problem 2.3

a-b.
training data exp confusion matrix:

0.5510	0.2430
0.0779	0.1280

training exp misclassification error = 0.3210
training data norm confusion matrix:

0.5250	0.1466
0.1039	0.2245

training norm misclassification error = 0.2505
testing exp confusion matrix:

0.6201	0.2009
0.0830	0.0961

testing exp misclassification error = 0.2838
testing exp sensitivity = 0.8820
testing exp specificity = 0.3235
testing norm confusion matrix:

0.5677	0.1092
0.1354	0.1878

testing norm misclassification error = 0.2445
testing norm sensitivity = 0.8075
testing norm specificity = 0.6324
c.
For the training data, the exponential prediction slightly underperformed the logistic regression model with 0.3210 misclassification, up from 0.3061. The normal distribution prediction overperformed the logistic regression model with 0.2505 misclassification, down from 0.3061.

For the testing data, the exponential prediction marginally underperformed the logistic regression model with 0.2838 misclassification, up from 0.2707. The normal distribution prediction also overperformed the logistic regression model with 0.2445 misclassification, down from 0.2707.

The best model in this assignment is the normal distribution model, with the lowest misclassification error of 0.2505 for training and 0.2445 for testing. Thise model also had a sensitivity of 0.8075 and a specificity of 0.6324.