

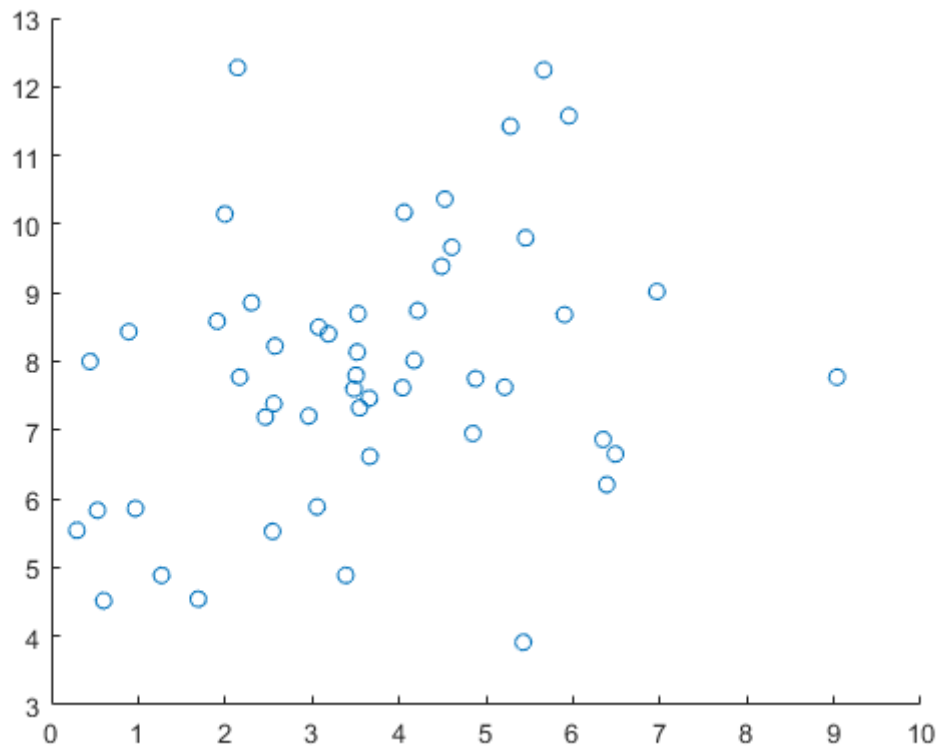
Homework 2

Adam Karl

February 18, 2021

1 Problem 1

a.



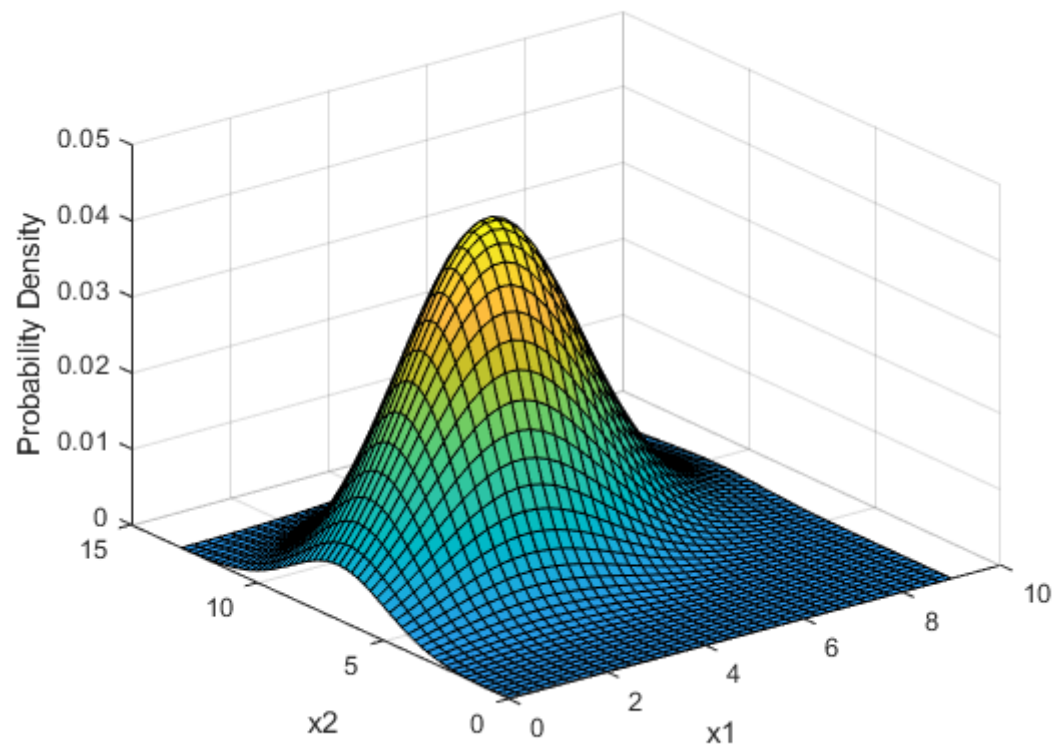
b.

mean vector:

$$\begin{pmatrix} 3.6377 & 7.8506 \end{pmatrix}$$

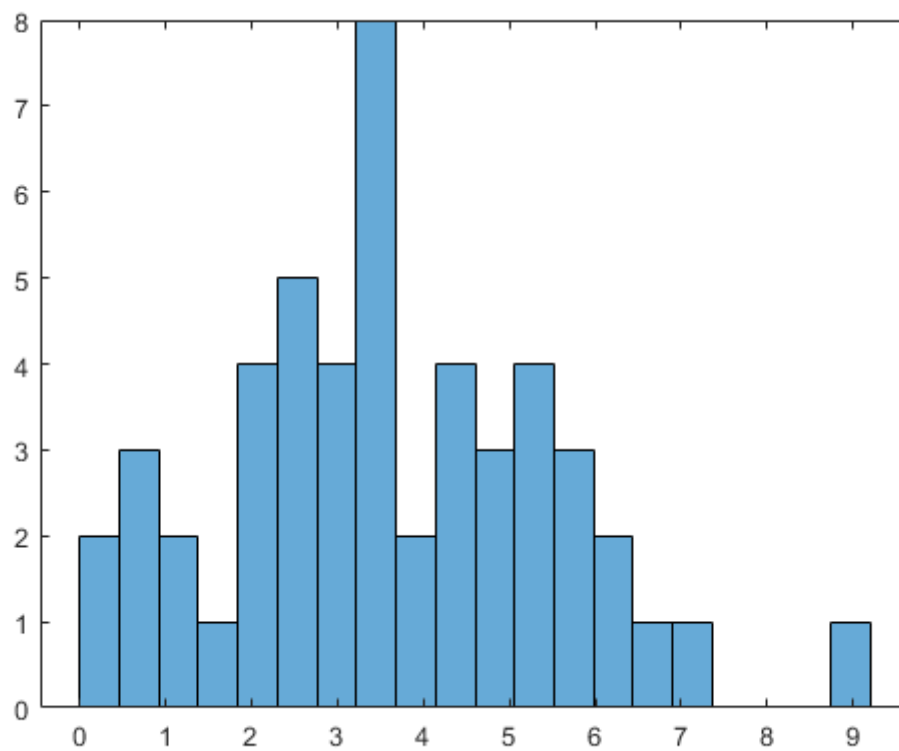
covariance matrix:

$$\begin{pmatrix} 3.6414 & 1.0779 \\ 1.0779 & 3.7831 \end{pmatrix}$$

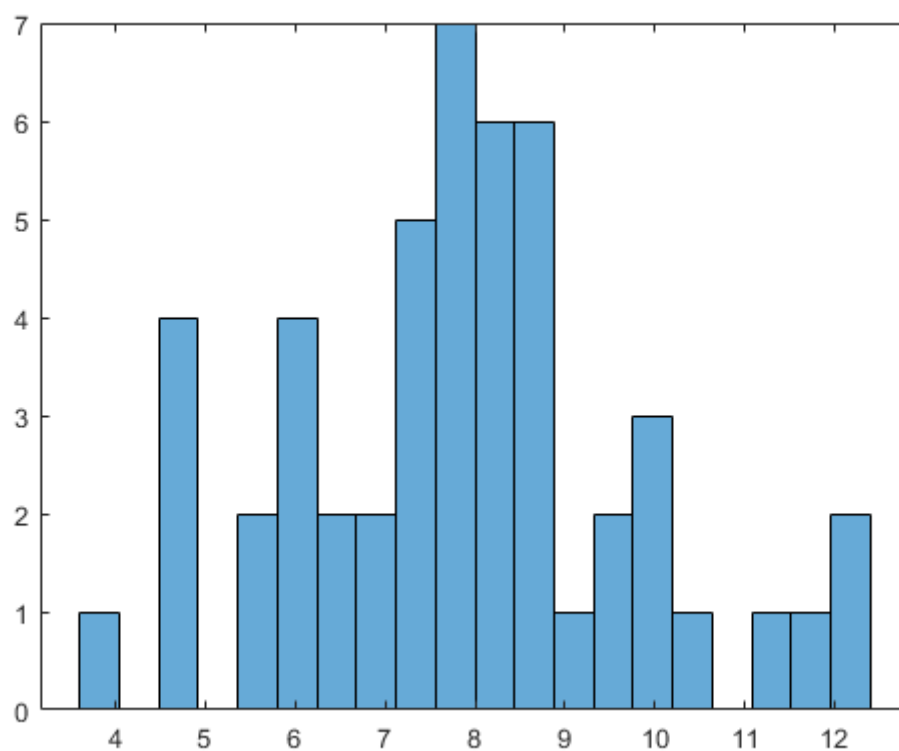


c.

- var 1
 - mean = 3.6377
 - variance = 3.6414
- var 2
 - mean = 7.8506
 - variance = 3.7831



variable 1



variable 2

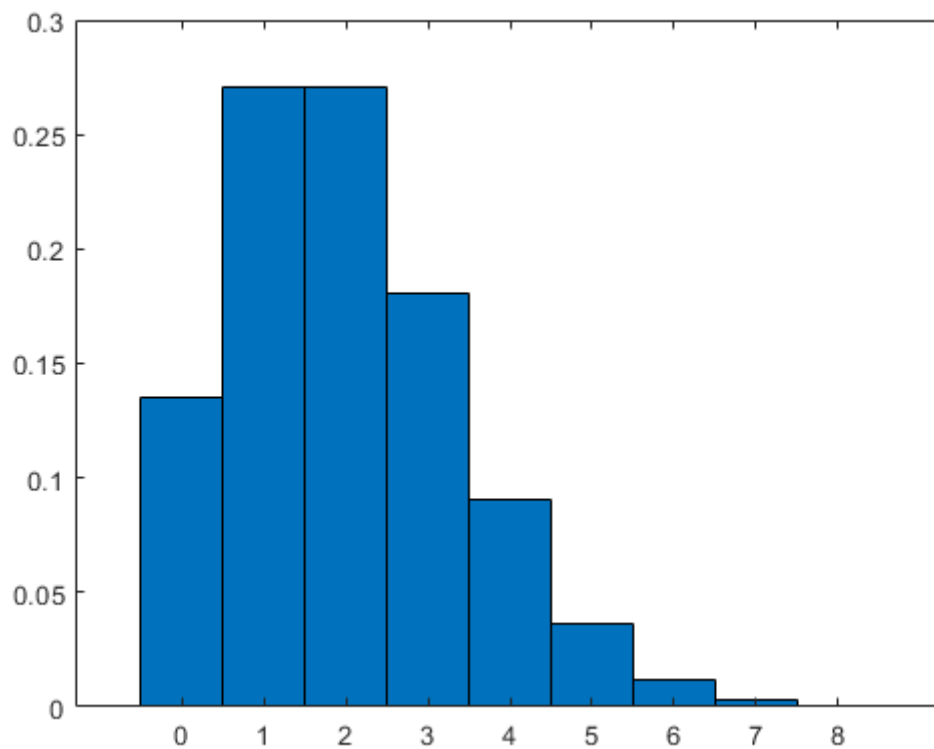
d.
The multivariate Gaussian model is certainly better than two separate univariate Gaussian models

because it includes correlation data. For this specific example, the correlation is extremely loose (value of 1.0779 from the covariance matrix) so it's not quite as important, but for other data sets it may have a much more noticeable correlation.

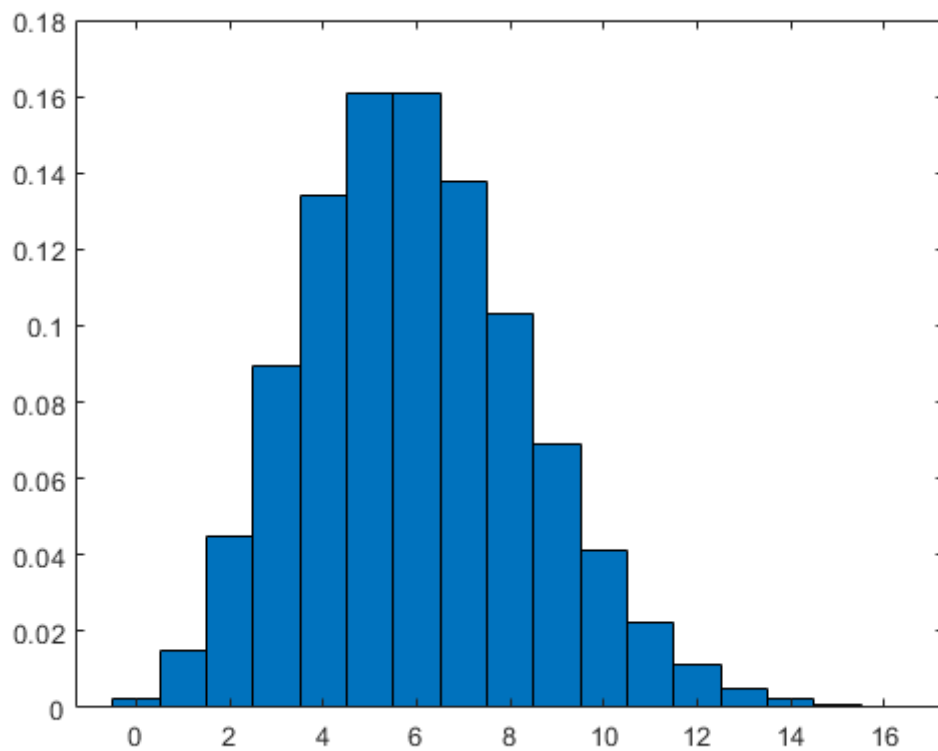
For instance, if var1 was father heights and var2 was children heights, then if you only looked at each variable individually then you would only see two normal-like distributions, and completely miss that taller fathers correspond to taller children.

2 Problem 2

a.



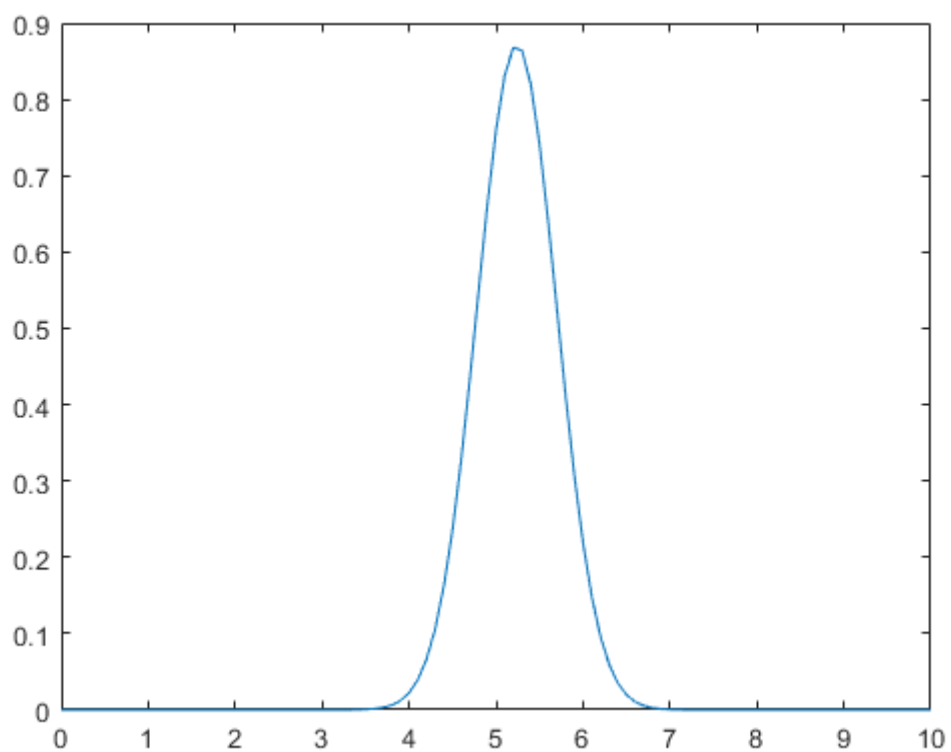
$\lambda = 2$



lambda = 6

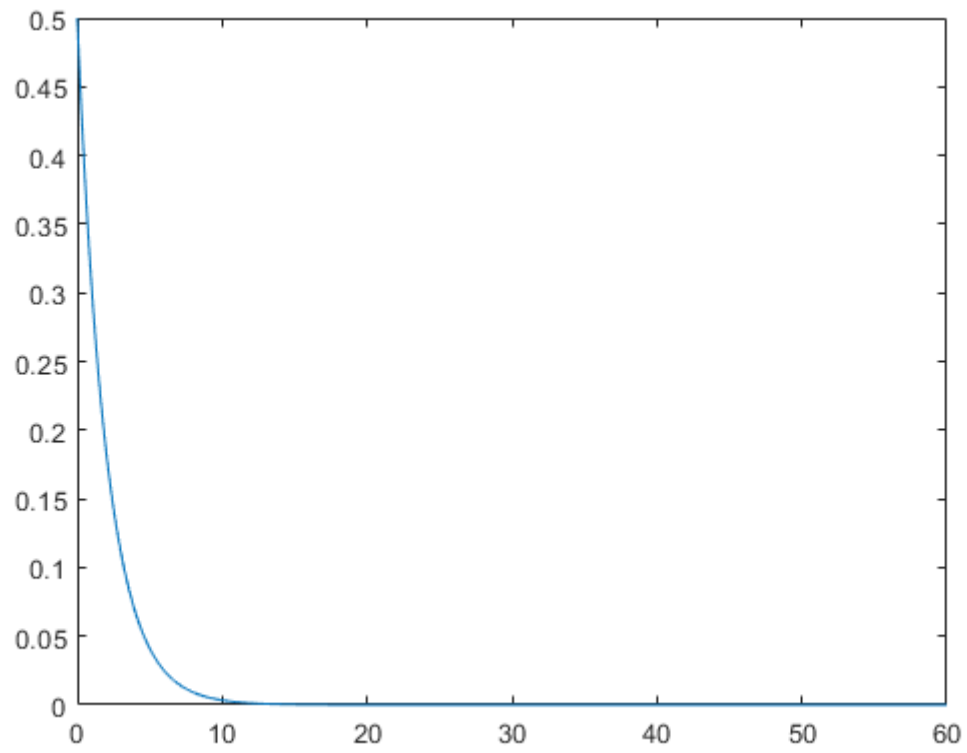
b. $\lambda_{ML} = \text{mean} = 5.2400$

Since variance = mean/n, we can calculate the standard deviation = 0.4578 and plot:

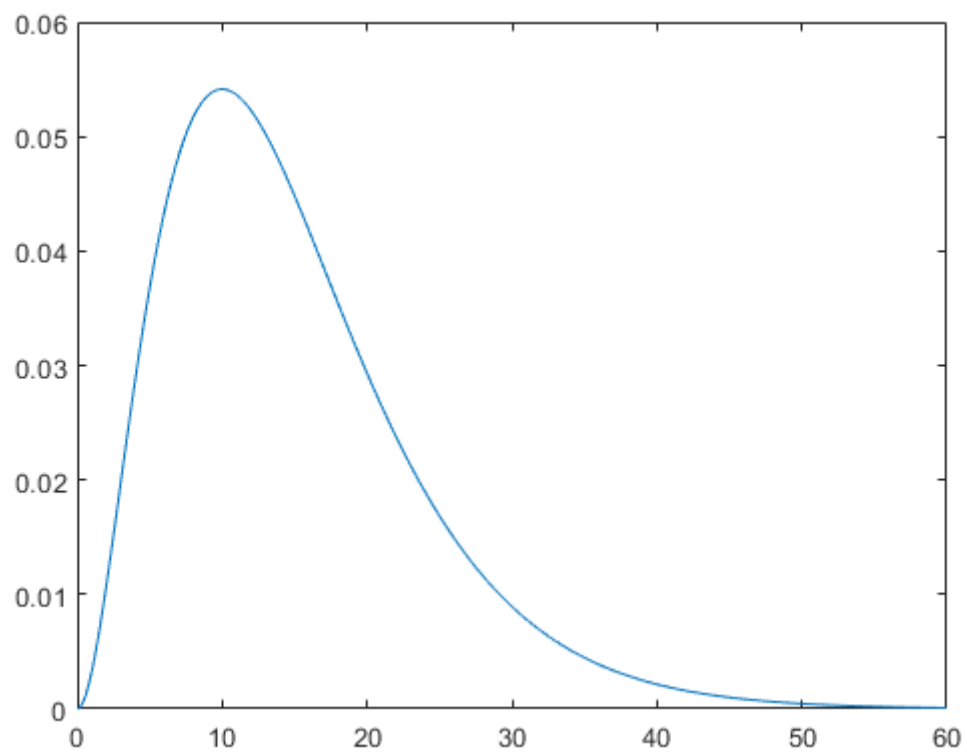


prob. function for ML parameter

c.



beta dist with $a=1$, $b=2$



beta dist with $a=3$, $b=5$

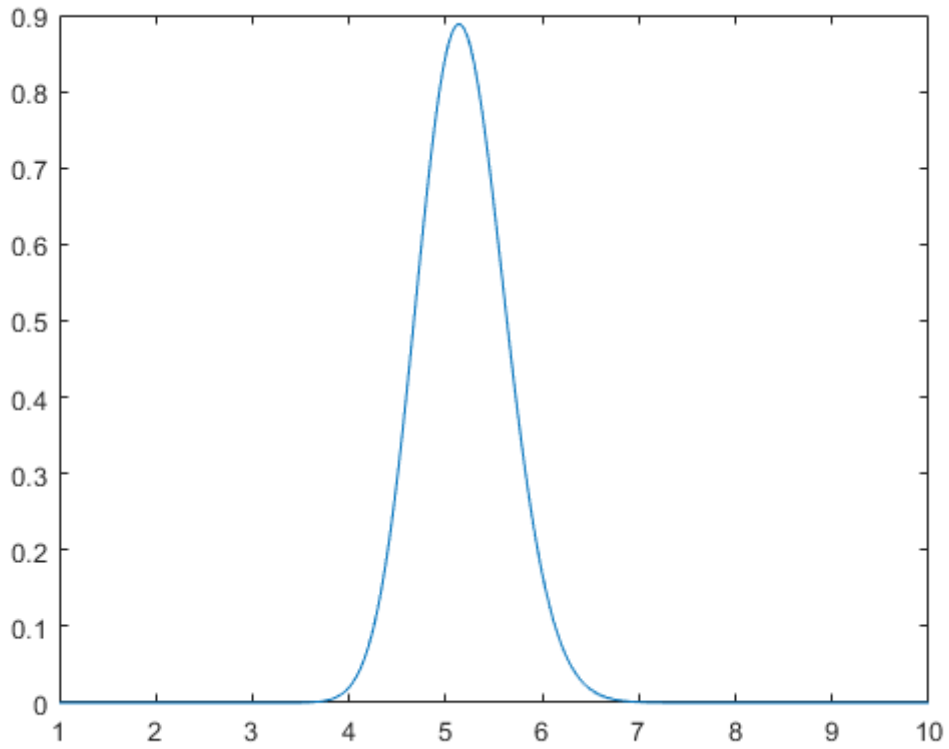
d.

a=1, b=2:

$$p(\lambda|D) \sim \text{Gamma}(\lambda|a + \sum_{i=1}^n x_i, \frac{b}{nb+1})$$

$$p(\lambda|D) \sim \text{Gamma}(1 + \sum_{i=1}^{25} x_i, \frac{2}{25(2)+1})$$

$$p(\lambda|D) \sim \text{Gamma}(132, \frac{2}{51})$$



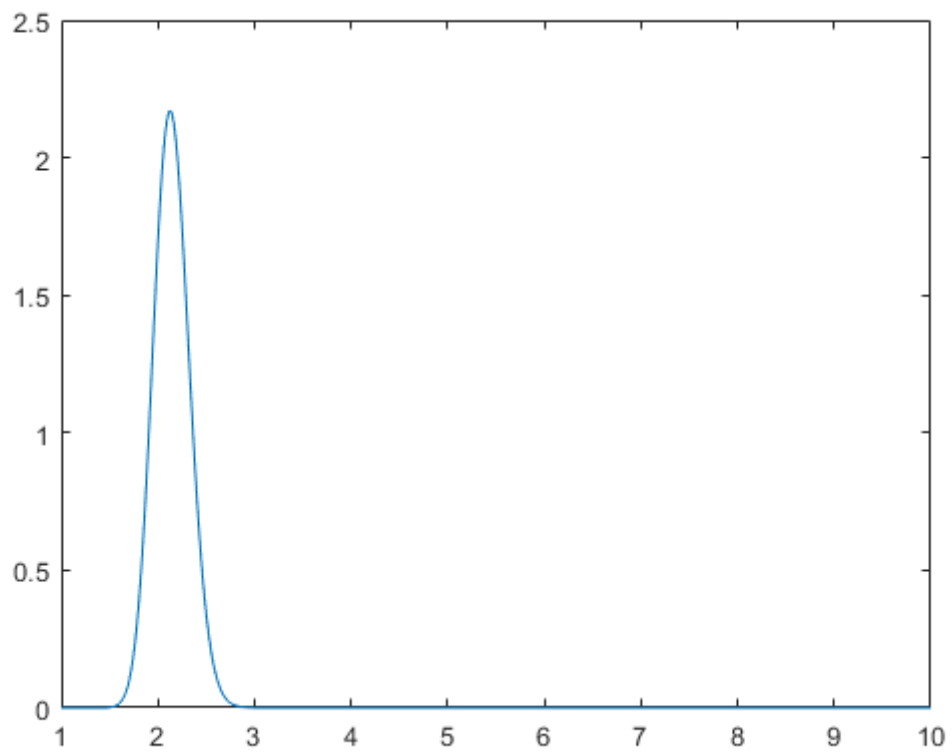
posterior 1

a=3, b=5:

$$p(\lambda|D) \sim \text{Gamma}(\lambda|a + \sum_{i=1}^n x_i, \frac{b}{nb+1})$$

$$p(\lambda|D) \sim \text{Gamma}(3 + \sum_{i=1}^{25} x_i, \frac{5}{25(5)+1})$$

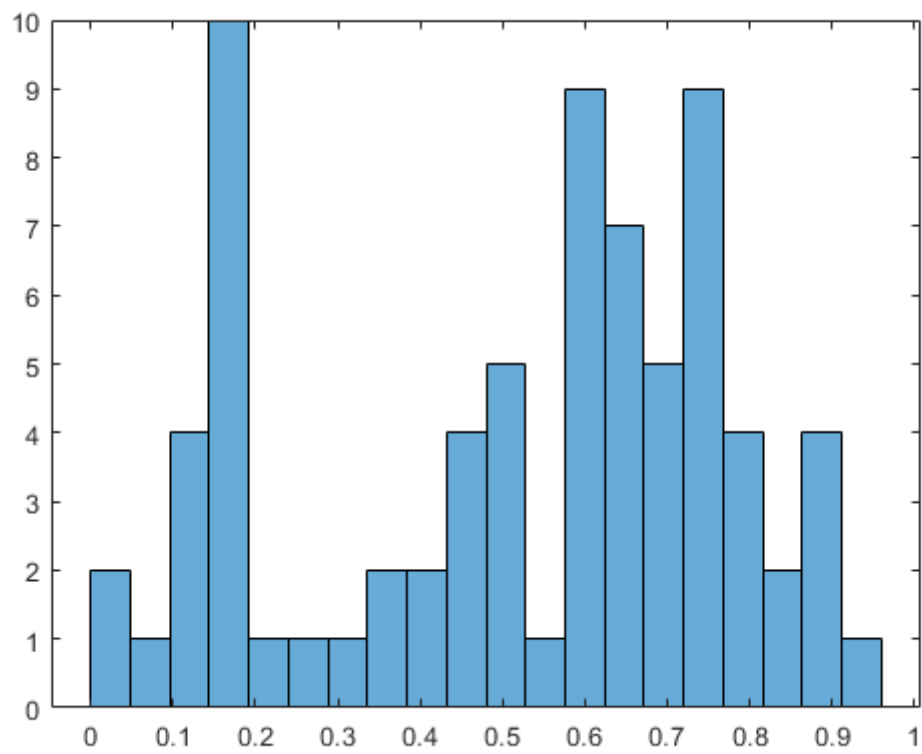
$$p(\lambda|D) \sim \text{Gamma}(135, \frac{1}{63})$$



posterior 2

3 Problem 3

Part 1



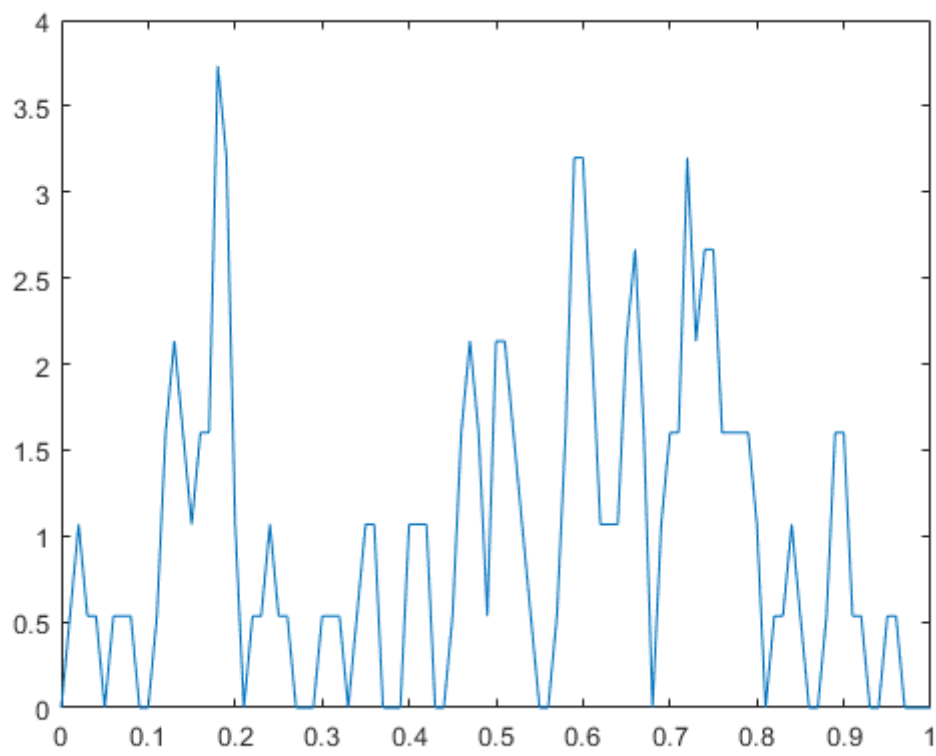
training data

Part 2

a.

with $h=0.025$

- $p(0.1362) = 1.6$
- $p(0.2928) = 0$
- $p(0.4491) = 1$
- $p(0.6517) = 2.1333$
- $p(0.7936) = 1.6$
- $p(0.6664) = 2.1333$

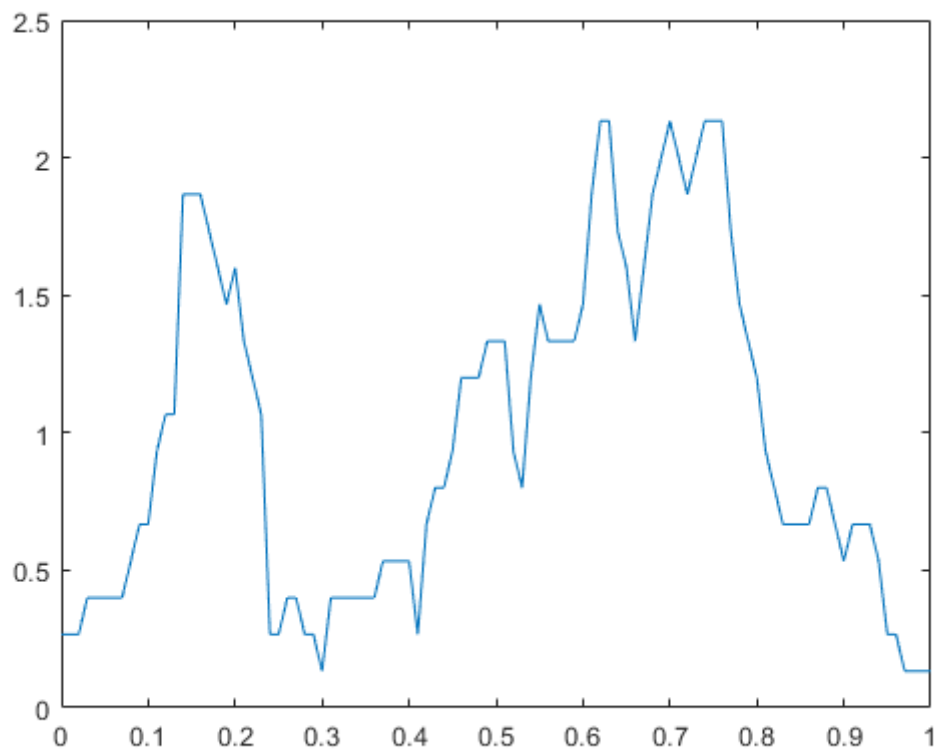


$h=0.025$

b.

with $h=0.1$:

- $p(0.1362) = 1.4667$
- $p(0.2928) = 0.2667$
- $p(0.4491) = 0.9333$
- $p(0.6517) = 1.6$
- $p(0.7936) = 1.3333$
- $p(0.6664) = 1.7333$



$h=0.1$

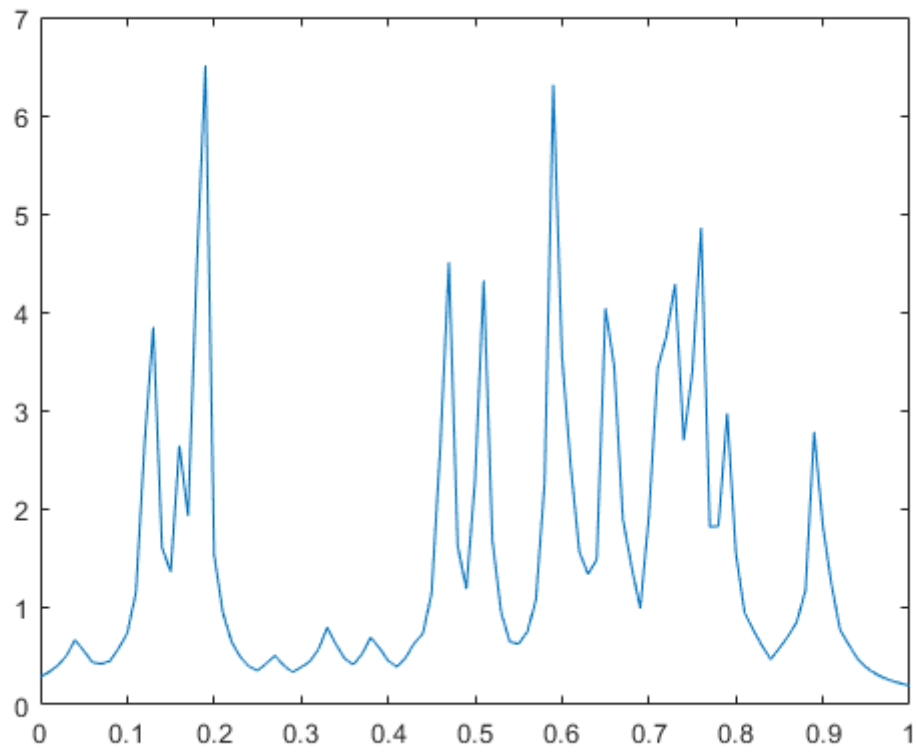
Increasing the h from 0.025 to 0.1 seems to smooth out the graph a lot, while decreasing the amplitude of the probability function. In the $h=0.025$ graph it's more obvious that individual points correspond to spikes in the graph. The peaks are still generally in the same places, but heavily smoothed out.

Part 3

a.

with $k=3$:

- $p(0.1362) = 1.7814$
- $p(0.2928) = 0.3448$
- $p(0.4491) = 1.0797$
- $p(0.6517) = 3.0180$
- $p(0.7936) = 3.0398$
- $p(0.6664) = 2.8816$

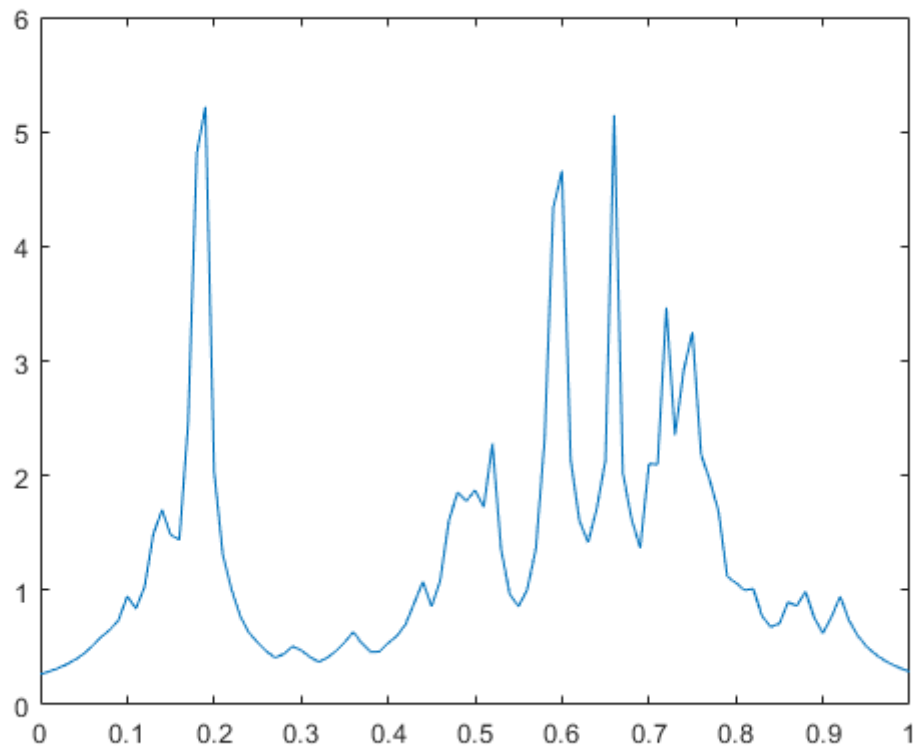


k=3

b.

with k=5

- $p(0.1362) = 1.8278$
- $p(0.2928) = 0.5311$
- $p(0.4491) = 0.8786$
- $p(0.6517) = 2.3900$
- $p(0.7936) = 0.9994$
- $p(0.6664) = 2.5931$



k=5

c.

Increasing the k from 3 to 5 seems to slightly smooth out the graph, while decreasing the amplitude of the probability function. The peaks are still generally in the same places, but are significantly decreased around $x=0.5$, supposedly because there were enough nearby points to satisfy $k=3$ easily, but the k NN function had to search much farther for a 5th point.