Homework 2

Adam Karl

February 11, 2021

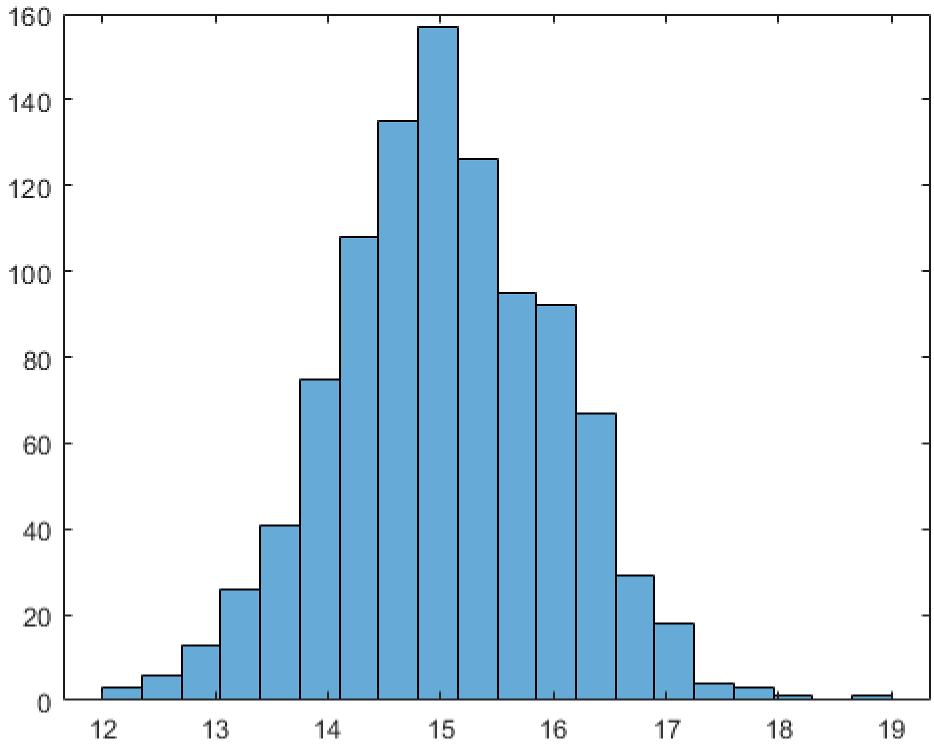
1 Problem 1

Part 1

- mean = 15.0415
- standard deviation = 5.0279

The calculated mean and standard deviations are each a few hundredths off of the true values. This is likely because matlab is rounding our input values to a certain number of decimal places.

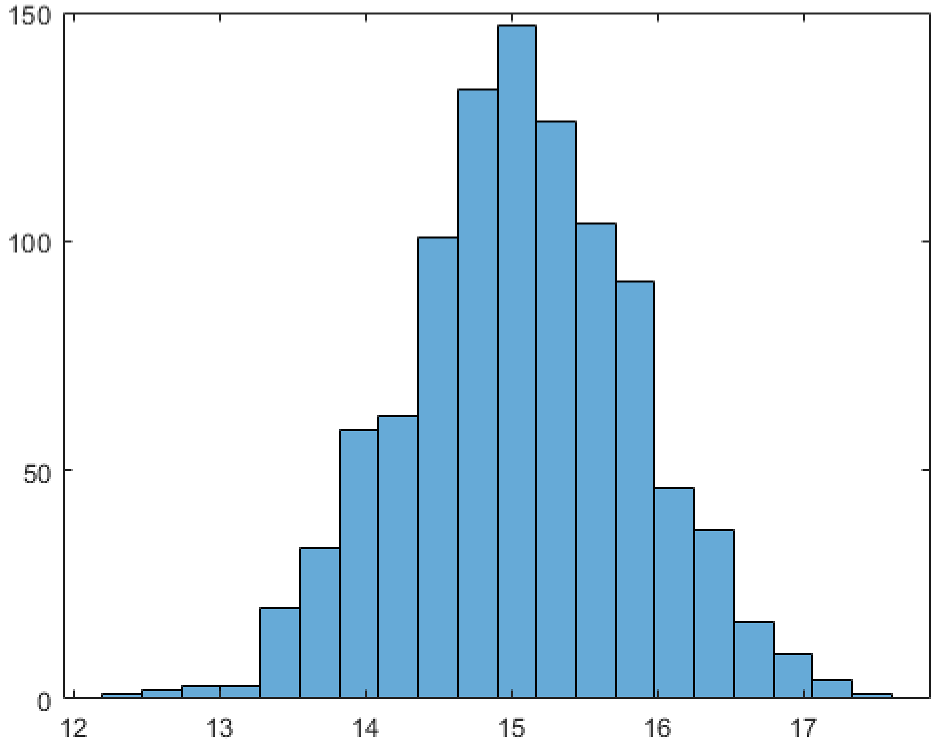
Part 4



Subsample size = 25

The means have a roughly normal distribution around the true mean of 15.
The mean of all subsample means was calculated to be 15.0587.

Part 5



Subsample size = 40

The means for a subsample size of 40 have a marginally tighter grouping than the 25-size subsample means around the true mean of 15. It's possible I was unlucky with my samples, but a subsample size of 40 does not appear to be much more accurate than a subsample size of 25.
The mean of all subsample means was calculated to be 15.0371.

Part 6
The first 25 elements have a mean of 14.5625.

The 95% confidence interval has a range of 12.6392 to 16.4974. The true mean of 15 is well within these bounds.

2 Problem 2

- m=1
  - test mean = 2.1346
  - test std = 3.2383
- m=2
  - test mean = 2.1077
  - test std = 3.2509
- m=3
  - test mean = 2.1798
  - test std = 3.2644
- m=4
  - test mean = 1.9794
  - test std = 3.2421

- m=5
  - test mean = 2.0685
  - test std = 3.2890
- m=6
  - test mean = 2.0369
  - test std = 3.3452
- m=7
  - test mean = 2.1467
  - test std = 3.3005
- m=8
  - test mean = 2.0701
  - test std = 3.3043
- m=9
  - test mean = 1.9808
  - test std = 3.2341
- m=10
  - test mean = 2.0516
  - test std = 3.1796

3 Problem 3

a.

- p(2) = 1/36
  - 1 way: (1,1)
- p(3) = 2/36 = 1/18
  - 2 ways: (1,2), (2,1)
- p(4) = 3/36 = 1/12
  - 3 ways: (1,3), (2,2), (3,1)
- p(5) = 4/36 = 1/9
  - 4 ways: (1,4), (2,3), (3,2), (4,1)
- p(6) = 5/36
  - 5 ways: (1,5), (2,4), (3,3), (4,2), (5,1)
- p(7) = 6/36 = 1/6
  - 6 ways: (1,6), (2,5), (3,4), (4,3), (5,2), (6,1)
- p(8) = 5/36
  - 5 ways: (2,6), (3,5), (4,4), (5,3), (6,2)
- p(9) = 4/36 = 1/9
  - 4 ways: (3,6), (4,5), (5,4), (6,3)

- p(10) = 3/36 = 1/12
  - 3 ways: (4,6), (5,5), (6,4)
- p(11) = 2/36 = 1/18
  - 2 ways: (5,6), (6,5)
- p(12) = 1/36
  - 1 way: (6,6)

b. EV $= 2(1/36) + 3(1/18) + 4(1/12) + 5(1/9) + 6(5/36) + 7(1/6) + 8(5/36) + 9(1/9) + 10(1/12) + 11(1/18) + 12(1/36) = 7$

c. The outcome of 4 has a $1/12$ chance. Therefore, the probability of NOT 4 is $1 - 1/12 = 11/12$. The probability of NOT 4, five consecutive times is $(11/12)^5 \approx 0.64722$.

The probability of an odd outcome is equal to p(3) + p(5) + p(7) + p(9) + p(11) $= 1/18 + 1/9 + 1/6 + 1/9 + 1/18 = 0.5$. The chance of an odd outcome all 5 trials is $0.5^5 = 0.03125$ (or 1/32).

4 Problem 4

Applying Bayes' theorem:

$$P(infected|positive) = \frac{P(positive|infected)P(infected)}{P(positive)}$$

$$P(infected|positive) = \frac{P(positive|infected)P(infected)}{P(falsepositive) + P(truepositive)}$$

$$P(infected|positive) = \frac{0.99 * 0.0001}{0.9999 * 0.01 + 0.0001 * 0.99}$$

$$P(infected|positive) = \frac{0.99 * 0.0001}{0.9999 * 0.01 + 0.0001 * 0.99}$$

$$P(infected|positive) \approx 0.0098$$

Since the probability that a person that tested positive is actually infected is only about 1%, I cannot recommend this test be widely adopted.

5 Problem 5

a. Simply integrate the probability density function over the range (a to b):

$$\int_a^b \frac{1}{b-a}\,dx$$

$$\frac{1}{b-a}\int_a^b 1\,dx$$

$$(\frac{1}{b-a})x\Big|_a^b$$

$$(\frac{1}{b-a})(b-a)$$

$$1$$

b. For a continuous distribution, to get the mean we integrate $x * f(x)$ over the entire range (here, from a to b):

$$\int_a^b x\frac{1}{b-a}\,dx$$

$$\frac{1}{b-a}\int_a^b x\,dx$$

$$(\frac{1}{b-a})\frac{1}{2}x^2\Big|_a^b$$

$$(\frac{1}{b-a})\frac{1}{2}(b^2-a^2)$$

$$(\frac{1}{b-a})\frac{1}{2}(b+a)(b-a)$$

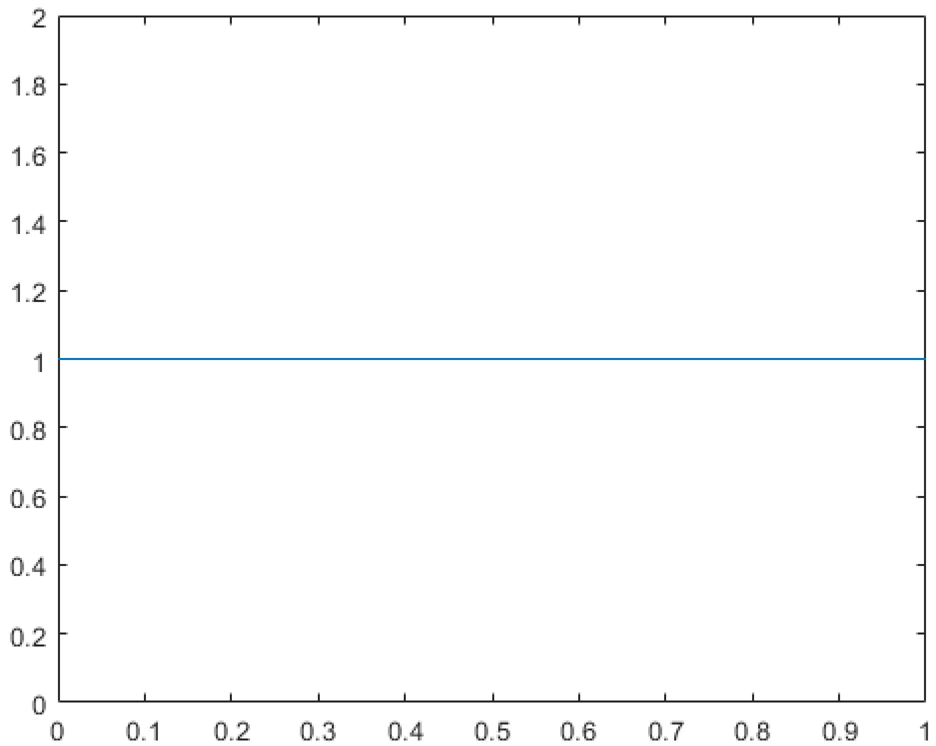
$$(\frac{1}{b-a})\frac{1}{2}(b+a)(b-a)$$

$$\frac{b+a}{2}$$

This makes logical sense, since it is the midpoint of the uniform distribution.

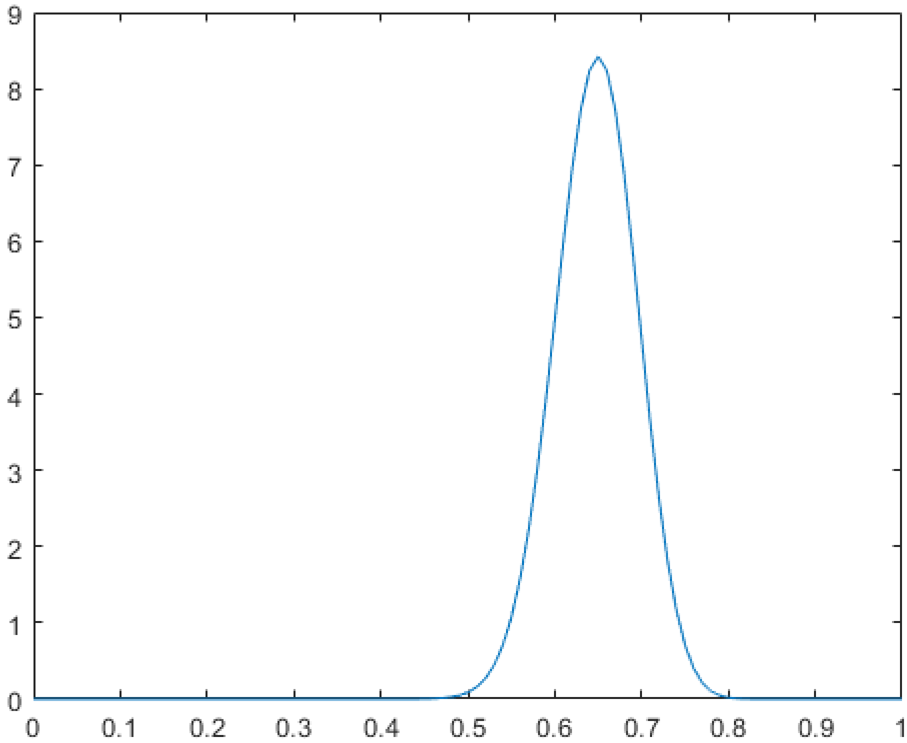
6 Problem 6

a. $65/100 = 0.65$
b.



prior

The prior distribution an a uniform distribution from 0 to 1.



posterior

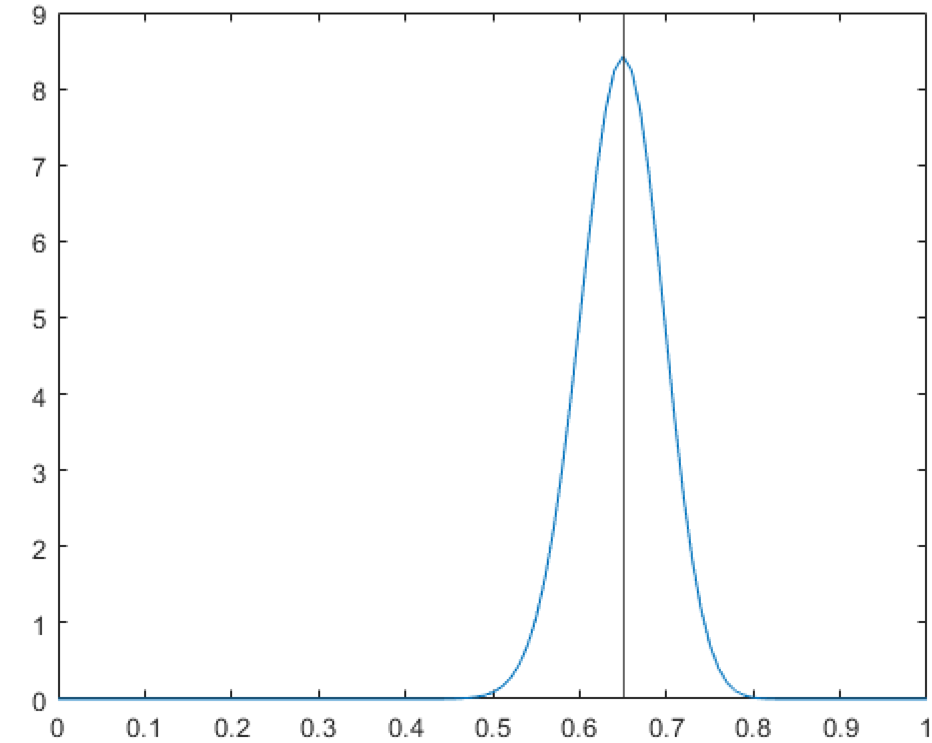
The posterior distribution is a normal-like distribution centered on our estimated theta of 0.65.

c.

$$\theta_{MAP} = \frac{N_1 + \alpha_1 - 1}{N_1 + N_2 + \alpha_1 + \alpha_2 - 2}$$

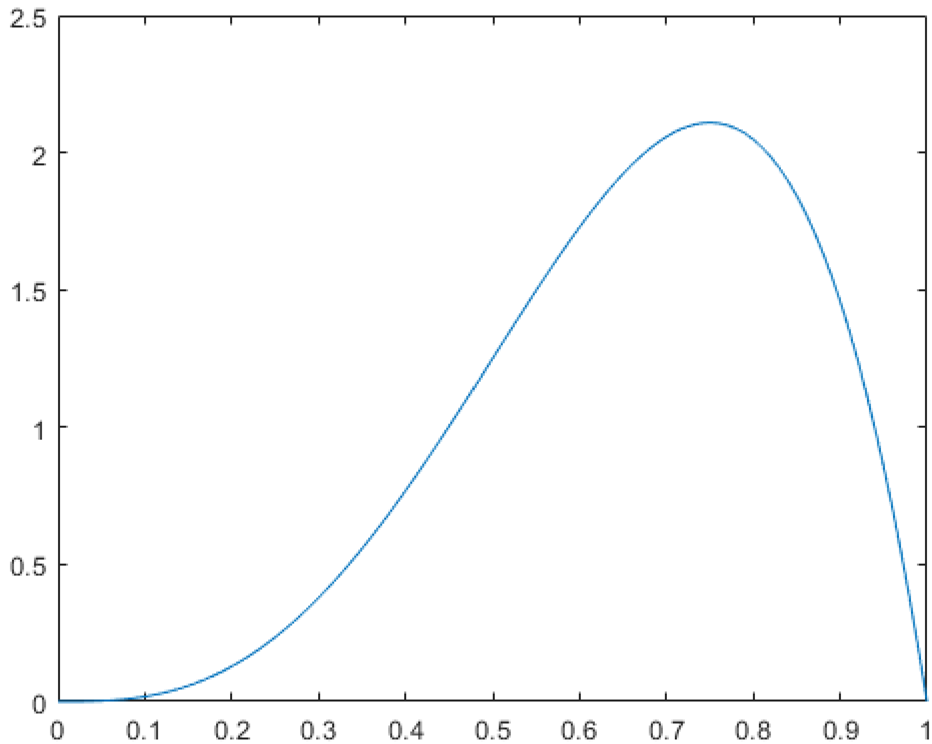
$$\theta_{MAP} = \frac{65 + 1 - 1}{65 + 35 + 1 + 1 - 2}$$

$$\theta_{MAP} = \frac{65}{100} = 0.65$$



$$\theta_{MAP} = 0.65$$

d.

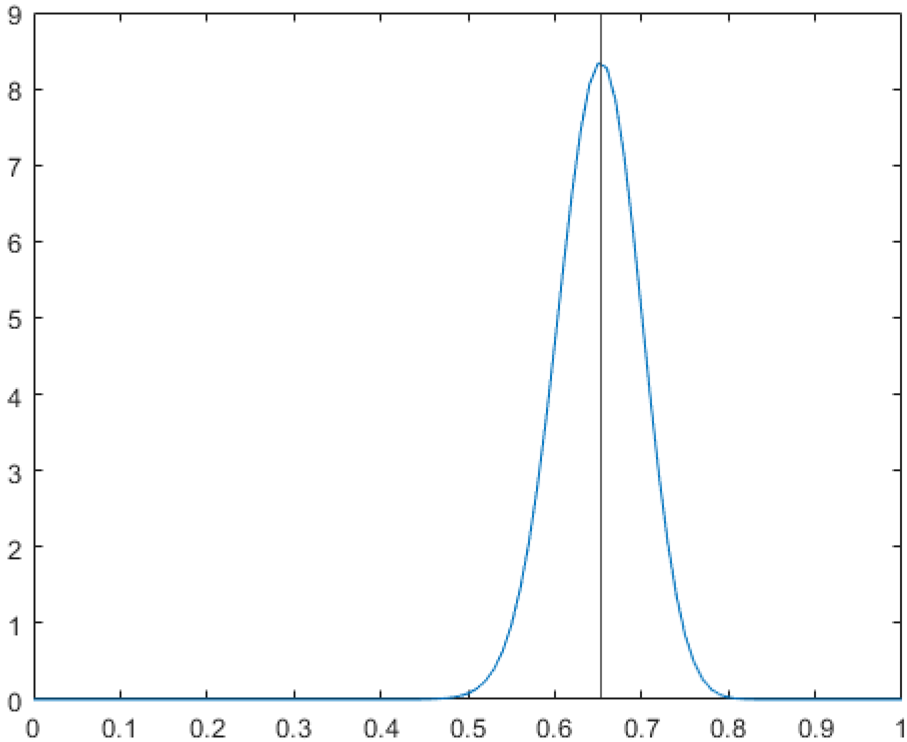


prior

$$\theta_{MAP} = \frac{N_1 + \alpha_1 - 1}{N_1 + N_2 + \alpha_1 + \alpha_2 - 2}$$

$$\theta_{MAP} = \frac{65 + 4 - 1}{65 + 35 + 4 + 2 - 2}$$

$$\theta_{MAP} = \frac{68}{104} \approx 0.6538$$



$\theta_{map} \approx 0.6538$