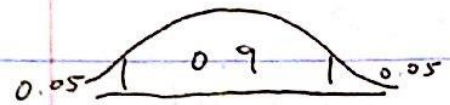
Homework 5.1

1) a) $\bar{x} = 447$ $\quad s^2 = 376.36$ $\quad n = 300$

$\alpha = 10\%$

$z = 1.645$



$$ME = 1.645\sqrt{\frac{376.36}{300}} = 1.8425$$

$$(445.158,\ 448.842) \text{ seconds}$$

b) $z = 2.58$

$$ME = 2.58\sqrt{\frac{376.36}{300}} = 2.885$$

$$(444.1,\ 449.9) \text{ seconds}$$

2) a) $$L(\theta) = \prod_{i=1}^{n} f(x_i|\theta) = \prod_{i=1}^{n} \lambda e^{-\lambda x_i}$$

$$LL(\theta) = \sum_{i=1}^{n} \log\left(\lambda e^{-\lambda x_i}\right)$$

$$= \sum_{i=1}^{n} \log\lambda - \lambda x_i$$

$$= n\log\lambda - \lambda\sum_{i=1}^{n} x_i$$

$$= n\log\lambda - \lambda Y$$

$$Y = \sum_{i=1}^{n} x_i$$

$$\frac{\partial LL(\lambda)}{\partial \lambda} = \frac{n}{\lambda} - Y = 0$$

$$\boxed{\lambda_{MLE} = \frac{n}{Y}} = \frac{n}{\sum_{i=1}^{n} X_i} = \frac{1}{\bar{X}}$$

b) Yes it is biased because the PDF for the exponential distribution is convex, so $E[f(x)] \geq f(E[x])$ and the bias is not 0.

c) $\lim_{n\to\infty} P(|\lambda_{MLE} - \lambda| < \varepsilon) = 1$ for $\varepsilon > 0$

$= \lim_{n\to\infty} P(|\frac{1}{\bar{x}} - \lambda| < \varepsilon) = 1$

Yes it is consistent because as n increases $\bar{x}$ approaches $\frac{1}{\lambda}$.

3) If the value of X_i is known, the values for all other X are known as well. The X ~~values~~ random variables are not independent, so it is problematic for the model learned by the Naive Bayesian Classifier.

4. No Laplace Estimators
   Class 0: tested 2, correctly classified 2
   Class 1: tested 2, correctly classified 2
   Overall: tested 4, correctly classified 4
   Accuracy: 1.000000

   With Laplace Estimators
   Class 0: tested 2, correctly classified 2
   Class 1: tested 2, correctly classified 2
   Overall: tested 4, correctly classified 4
   Accuracy: 1.000000

5. No Laplace Estimators
   Class 0: tested 48, correctly classified 52
   Class 1: tested 76, correctly classified 83
   Overall: tested 124, correctly classified 135
   Accuracy: 0.918519

   With Laplace Estimators
   Class 0: tested 48, correctly classified 52
   Class 1: tested 76, correctly classified 83
   Overall: tested 124, correctly classified 135
   Accuracy: 0.918519

6. No Laplace Estimators
   Class 0: tested 10, correctly classified 15
   Class 1: tested 135, correctly classified 172
   Overall: tested 145, correctly classified 187
   Accuracy: 0.775401

   With Laplace Estimators
   Class 0: tested 10, correctly classified 15
   Class 1: tested 130, correctly classified 172
   Overall: tested 140, correctly classified 187
   Accuracy: 0.748663

7. Laplace Estimators did not change the accuracy of the vote data, but decreased the accuracy for the heart data. In general, Laplace Estimators improve accuracy when test data produce unrepresentative MLE estimates, such as an MLE estimate of 0 for an event that can actually occur. Otherwise, Laplace Estimators are not better than MLEs for classification accuracy.