MET CS 555 Assignment 6 – 20 points

SUBMISSION REQUIREMENTS: **Please submit a single document (word or PDF) for submission.  Your submission should contain a summary of your results (and answers to questions asked on the homework) as well as your R code used to generate your results (please append to the end of your submission). Please use R for the calculations whenever possible. You will lose points if you are not utilizing R.**

**The data “A06.csv” consists of body temperature and heart rate measurements for 65 men and 65 women.**  **Save the data to a file, and read the data file into R. Use this data to address the following questions. Data (1=males, 2 =females)**

1. We are interested in whether the proportion of men and women with body temperatures greater than or equal to 98.6 degrees Fahrenheit are equal. Therefore, we need to dichotomize the body temperature variable. Create a new variable, called **“temp\_level”** in which **temp\_level = 1** if body temperature **>= 98.6 and temp\_level=0** if body temperature **< 98.6**. **(2 points)**
2. Summarize the data relating to body temperature level (i.e., the variable you created above) by sex. **(2 points)**
3. Calculate the risk difference for high body temperature levels between men and women. Formally test (at the **alpha=0.05** level) whether the proportion of people with higher body temperatures (greater than or equal to **98.6**) is the same across men and women based on this effect measure. You should show all five steps in the 5-step recipe for testing. **(4 points)**
4. Perform a logistic regression with sex as the only explanatory variable. Formally test (at the **alpha=0.05** level) if the odds of having a temperature **greater than or equal to 98.6** is the same between males and females. Again, please show all five steps. Additionally, include the odds ratio for sex and the associated **95%** confidence interval in your summary, and interpret the value of the odds ratio. Lastly, what is the c-statistic for this model? **(5 points)**
5. Perform multiple logistic regression predicting body temperature level from sex and heart rate. Briefly summarize the output from this mode (no need to go through all five steps). Give the odds ratio for sex. Also, report the odds ratio for heart rate **(for a 10-beat increase)**. What is the c-statistic of this model? **(5 points)**
6. Which model fits the data better? Support your response with evidence from your output. Present the ROC curve for the model you choose. (**2 points)**