

STT 811

In-Class Assignment 7

This problem will use the Heart dataset. The target is the AHD field.

1. Create a target based on the AHD field with numerical values of 0 and 1.
 - a. `heart$y <- ifelse(heart$AHD=="Yes",1,0)`
2. Build a logistic regression model for your target based on MaxHR, RestBP, and ChestPain. How significant are the coefficients?
 - a. `heart_mod <- glm(data = heart, y ~ MaxHR + RestBP + as.factor(ChestPain), family = binomial)`
 - b. all values are significant to at least the 0.05 level, some even further
3. Calculate the risk of heart disease based on this model for someone with a Maximum Heart Rate of 170, a Resting Blood Pressure of 145, and nontypical Chest Pain. Express the risk both as a probability and as an odds ratio.
 - a. $b_0 = 2.915699$, $b_1 = -0.032991$, $b_2 = 0.021750$, $b_3 = -2.010243$
 - b. $x_1 = 170$, $x_2 = 145$
 - c. $y_1 <- 1/(1+\exp(-(b_0+b_1*x_1+b_2*x_2+b_3)))$
 - d. $y_1 = 0.1751926$
4. Create the confusion matrix for this model, based on the y values and the fitted values.
 - a. `confusionMatrix(data = as.factor(as.integer(2*heart_mod$fitted.values)), reference = as.factor(heart$y))`

```
Confusion Matrix and Statistics

          Reference
Prediction  0    1
          0 127  33
          1  37 106

              Accuracy : 0.769
              95% CI   : (0.7174, 0.8152)
    No Information Rate : 0.5413
    P-Value [Acc > NIR] : <2e-16

              Kappa   : 0.5358

  Mcnemar's Test P-Value : 0.7199

              Sensitivity : 0.7744
              Specificity : 0.7626
              Pos Pred Value : 0.7938
              Neg Pred Value : 0.7413
              Prevalence : 0.5413
              Detection Rate : 0.4191
              Detection Prevalence : 0.5281
              Balanced Accuracy : 0.7685

              'Positive' Class : 0
```

b.