Problem Set 8

Cody Frisby

April 7, 2016

Part I

1) Prelim Analysis

Here are the observed and expected values between y and t from our data set

	0	1	Predicted 0	Predicted 1
Observed 0	4	14	6.685714	11.31429
Observed 1	9	8	6.314286	10.68571

where the rows are the t values and the columns are the y values. And the test statistic from the χ^2 test is 3.5337538 and the p value = 0.0601315. It appears there is some effect on sore throats from device but not statistically significant at the $\alpha = 0.05$ level. Sore throat rate for mask = 0.7777778, and the sore throat rate for tube = 0.4705882.

2) Logistic Model

- The fitted model is
$$ln(\frac{\pi}{1-\pi}) = -2.2135821 + 0.0703828duration$$
and the p value for duration is 0.0083106.

- $e^{0.0703828(10)}$ would give us the odds ratio of waking up with a sore throat for a duration lasting ten additional minutes. $e^{0.0703828(10)} = 2.0214757$

- The predicted probabilities for waking up with a sore throat given duration is 15 vs 45 minutes is

15 minutes	0.2390599
45 minutes	0.7218472

- The AIC for our model is 37.6513406. The BIC is 40.7620367.
- The p-value from the Hoslem-Lemeshow (GOF) test is 0.2305743. This p value indicates that our model is "good", meaning that the model fits reasonably well. Here I specified 8 bins. I display the resulting matrix for illustration.

bin	y0	y1	yhat0	yhat1
[0.239,0.256]	5	0	3.8047	1.1953
(0.256,0.388]	3	3	3.9085	2.0915
(0.388,0.474]	1	2	1.5765	1.4235
(0.474,0.646]	1	4	1.9376	3.0624
(0.646,0.738]	1	2	0.8345	2.1655
(0.738,0.901]	2	2	0.5608	3.4392

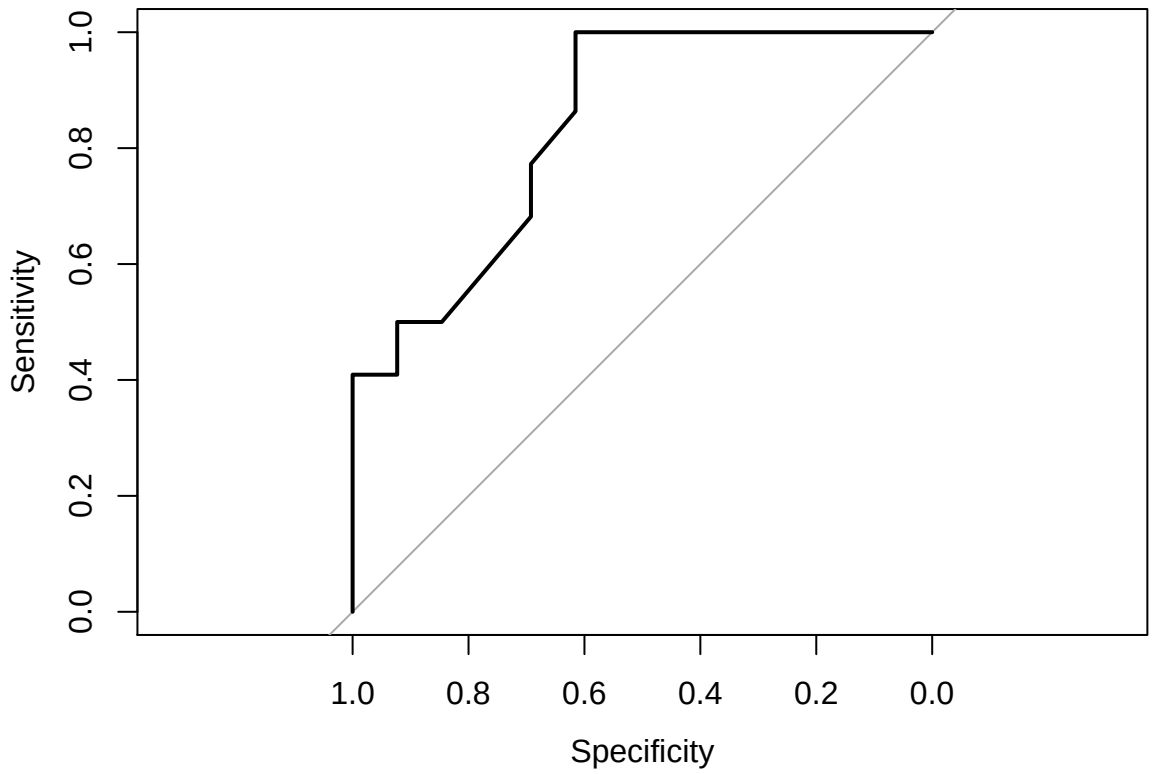
bin	y0	y1	yhat0	yhat1
(0.901,0.955]	0	5	0.3237	4.6763
(0.955,0.999]	0	4	0.0538	3.9462

- I had to use a custom function to have R compute the percent concordant for this problem. The value is 82.8671329.
- For the contingency table I used $\pi > 0.5$ as my cutoff threshold which results in

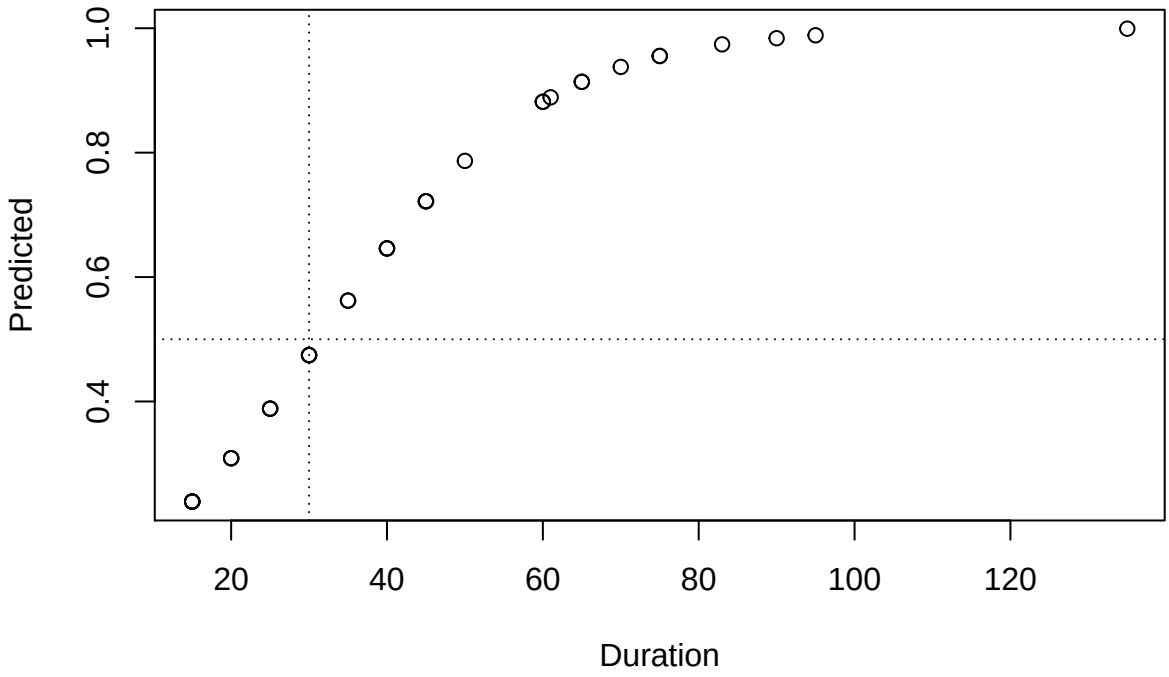
	Predicted0	Predicted1
Observed 0	9	4
Observed 1	5	17

so the sensitivity of our model, $P(\hat{y} = 1 \mid y = 1)$, is 0.7727273. And the specificity of our model, $P(\hat{y} = 0 \mid y = 0)$, is 0.6923077. And the overall classification rate is 0.7428571.

- The ROC curve



- Plot of predicted vs. duration. I've added a horizontal line at 0.5 and a vertical line at 30 to help with the visual inspection.



It appears that the 50-50 point is at around duration = 30, perhaps slightly longer than 30.

- It would appear that duration of surgery has an impact on the probability of waking up with a sore throat. It is statistically significant and the confidence interval of the odds ratio of duration does not contain one, [1.0267514, 1.1427819].

Part II

1) Mathematical Model

- Model $log(\frac{\pi}{1-\pi}) = \beta_0 + \beta_1 duration + \beta_2 type$ where β_1 is the expected change in log(odds) per unit change in duration and β_2 is the expected change in log(odds) when the tracheal tube is used instead of mask, and whereπ is equal to the probability of getting a sore throat.

2) Fit Model

- The fitted model is

$$ln(\frac{\pi}{1-\pi}) = -1.4173413 + 0.0686778 duration - 1.6589491 type$$

  and the p value for **duration** is 0.0093135 and for **type**, 0.0722359.

- For a surgery lasting an additional 10 minutes, type of device held constant, we expect approximately the same increase in the odds as before of waking up with a sore throat, 1.9873018.

- Duration held constant, the expected change in odds of waking up with a sore throat when a tracheal tube is used is 0.1903389 times 100, percent that of when the mask is used. Our odds of getting a sore throat are actually slightly lower when the tube is used then when the mask is used. For examples, let's say the duration is 10 minutes, comparing the two equations for tube and mask we would have

$$odds = e^{-1.4173413+0.0686778(10)-1.6589491}$$

and

$$odds = e^{-1.4173413+0.0686778(10)}$$

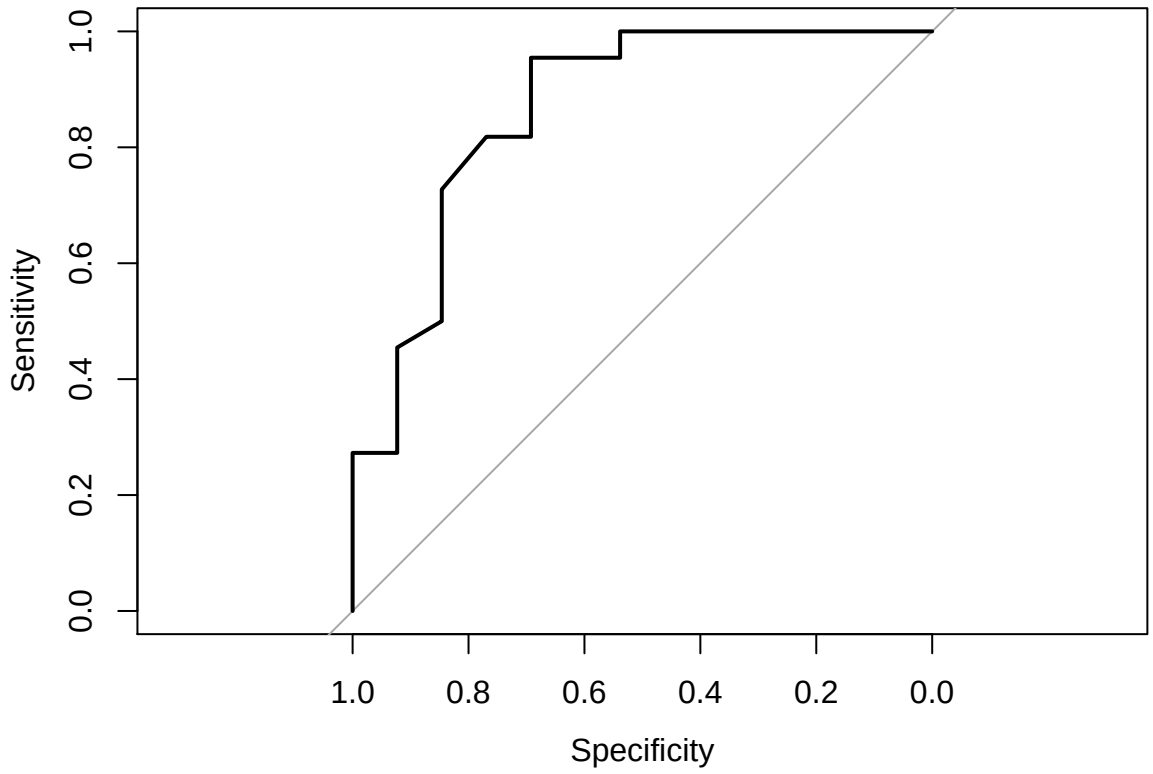
resulting in 0.0916744 when tube is used and 0.4816375 when mask is used.

- Predicted probabilities for d = 15 and t = 1 yields 0.1144447 and for d = 45 an t = 0, 0.8419965.
- AIC = 36.1379403 and BIC = 40.8039845.
- P value for the Hosmer-Lemeshow test is 0.4742036. We do not reject the null hypothesis. The model fits reasonably well.
- Percent concordance is 86.013986.
- For the contingency table I used $\pi > 0.5$ as my cutoff threshold which results in

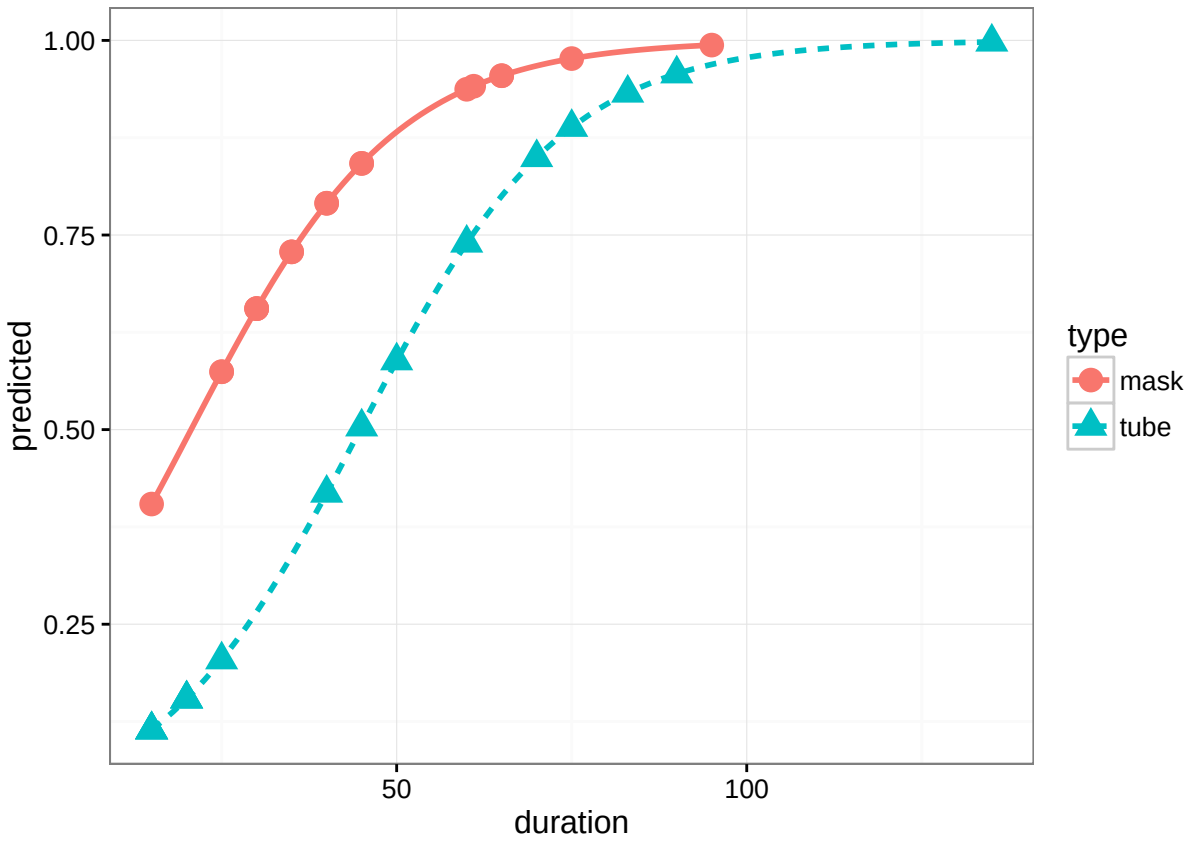
	Predicted0	Predicted1
Observed 0	9	4
Observed 1	1	21

so the sensitivity of our model, $P(\hat{y} = 1 \mid y = 1)$, is 0.9545455. And the specificity of our model, $P(\hat{y} = 0 \mid y = 0)$, is 0.6923077. And the overall classification rate is 0.8571429.

- ROC curve



- Plot of predicted vs duration by device type.



Eyeballing this plot, it appears that the 50-50 point is around 20 minutes when the mask is used and just under 50 minutes when the tube is used.

3) Liklihood Ratio Test

The resulting test statistic from this test is 3.5134003 and the p value is 0.0608744. This is borderline significant, and we should consider other tests/statistics when considering which model to use.

4) Which Model?

	AIC	BIC	Concordant	Sensitivity	Specificity	Classification
RM	37.65134	40.76204	82.86713	0.7727273	0.6923077	0.7428571
FM	36.13794	40.80398	86.01399	0.9545455	0.6923077	0.8571429

Considering this table, I would choose the full model. We gain a whole lot of *sensitivity* and are only slightly worse on BIC. AIC and concordance are also more desirable.

5) Conclusions

The answer to the initial research question is yes! Duration of the surgery does have an affect on whether patients will experience a sore throat when waking up, with the longer the duration the more likely. Additionally, knowing which device is used, with tube being better choice (not considering interactions), we

can increase the sensitivity of the model from 0.7727273 to 0.9545455. Our initial look at the type of device, running a simple χ^2 test it was borderline at the $\alpha = 0.05$ level. But considering all things, I would choose to include device type in our model as it does affect prediction.

R code:

```
sore <- xlsx::read.xlsx("~/Documents/MATH3710/ProblemSets/problem8/sore throat.xlsx", sheetIndex = 1)
names(sore) <- c("d", "t", "y")
# y is the response, 1 = yes, patient had sore throat
# d = duration of surgery
# t = type of device to secure the airway, 1 = tracheal tube, 0 = laryngeal mask
# T is built into R as TRUE, so I changed T to t.
test <- with(sore, chisq.test(t, y, correct = FALSE))
fit <- glm(y ~ d, data = sore, family = binomial(link = "logit"))
b <- coef(fit)
p <- summary(fit)$coef[,4]
s.10 <- exp(coef(fit)) * 10
pred.odds <- exp(predict(fit, data.frame(d = c(15,45))))
probs <- pred.odds/(1 + pred.odds)
names(probs) <- c("15 minutes", "45 minutes")
# to run Hosmer and Lemeshow tests, we install the package "ResourceSelcetion"
library(ResourceSelection)
h.test <- hoslem.test(sore$y, fit$fitted.values, g = 8)
h <- cbind(round(h.test$observed,4), round(h.test$expected,4))
bin <- rownames(h)
rownames(h) <- NULL
h <- cbind(bin, h)
#############################################################
# Function OptimisedConc : for concordance, discordance, ties
# The function returns Concordance, discordance, and ties
# by taking a glm binomial model result as input.
# Although it still uses two-for loops, it optimises the code
# by creating initial zero matrices
#############################################################
OptimisedConc=function(model)
{
  Data = cbind(model$y, model$fitted.values)
  ones = Data[Data[,1] == 1,]
  zeros = Data[Data[,1] == 0,]
  conc=matrix(0, dim(zeros)[1], dim(ones)[1])
  disc=matrix(0, dim(zeros)[1], dim(ones)[1])
  ties=matrix(0, dim(zeros)[1], dim(ones)[1])
  for (j in 1:dim(zeros)[1])
  {
    for (i in 1:dim(ones)[1])
    {
      if (ones[i,2]>zeros[j,2])
      {conc[j,i]=1}
      else if (ones[i,2]<zeros[j,2])
      {disc[j,i]=1}
      else if (ones[i,2]==zeros[j,2])
      {ties[j,i]=1}
```

```
    }
  }
  Pairs=dim(zeros)[1]*dim(ones)[1]
  PercentConcordance=(sum(conc)/Pairs)*100
  PercentDiscordance=(sum(disc)/Pairs)*100
  PercentTied=(sum(ties)/Pairs)*100
  return(list("Percent Concordance"=PercentConcordance,"Percent Discordance"=PercentDiscordance,"Percen
}
o <- OptimisedConc(fit)
theprobs <- fit$fitted.values
t <- as.matrix(table(fit$y, 1*(theprobs > .5)))
# how often are we right?
sensitivity <- t[2,2]/sum(t[2,])
specificity <- t[1,1]/sum(t[1,])
overall <- (sum(diag(t)))/sum(t)
obs <- rownames(t)
rownames(t) <- NULL
t <- cbind(obs, t)
library(pROC)
r <- roc(fit$y ~ fit$fitted.values)
fig <- plot(roc(fit$y, fit$fitted.values))
plot(sore$d, fit$fitted.values, ylab="Predicted", xlab="Duration")
abline(h=0.5, lty = 3); abline(v=30, lty = 3)
fit2 <- glm(y ~ d+t, data = sore, family = binomial(link = "logit"))
b2 <- coef(fit2)
p <- summary(fit2)$coef[,4]
s.2 <- exp(coef(fit2)) * 10
odds <- exp(coef(fit2))
pred2 <- predict(fit2, data.frame(d = 10, t = c(1,0)))
pred2.odds <- exp(pred2)
# note, t = 1 is tracheal tube, t = 0 is mask.
pred3 <- predict(fit2, data.frame(d = c(15, 45), t = c(1,0)))
pred3.odds <- exp(pred2)
p3 <- pred3.odds/(1+ pred3.odds)
h.test2 <- hoslem.test(sore$y, fit2$fitted.values, g = 8)
h <- cbind(round(h.test2$observed,4), round(h.test2$expected,4))
bin <- rownames(h)
rownames(h) <- NULL
h <- cbind(bin, h)
o2 <- OptimisedConc(fit2)
theprobs2 <- fit2$fitted.values
t2 <- as.matrix(table(fit2$y, 1*(theprobs2 > .5)))
# how often are we right?
sensitivity2 <- t2[2,2]/sum(t2[2,])
specificity2 <- t2[1,1]/sum(t2[1,])
overall2 <- (sum(diag(t2)))/sum(t2)
obs <- rownames(t2)
rownames(t2) <- NULL
t2 <- cbind(obs, t2)
library(pROC)
r <- roc(fit2$y ~ fit2$fitted.values)
fig <- plot(roc(fit2$y, fit2$fitted.values))
# adding a new column to sore
```

```r
sore$type <- 0
sore[sore$t == 1, "type"] <- "tube"
sore[sore$t == 0, "type"] <- "mask"
sore$type <- as.factor(sore$type)
sore$predicted <- fit2$fitted.values
library(ggplot2) # using ggplot2 to create this plot
g <- ggplot(data = sore, aes(x = d, y = predicted)) + xlab("duration")
g <- g + geom_point(aes(shape = type, colour = type), size = 4) + theme_bw()
g <- g + geom_smooth(aes(group=type, linetype = type, color = type),
                     method = "glm",
          method.args = list(family = "binomial"), se = F)
g
lrt <- anova(fit, fit2, test = "LRT")
s.lrt <- lrt[[4]][2]
p.lrt <- lrt[["Pr(>Chi)"]][2]
rm.concord <- o[1]; fm.concord <- o2[1]
rm <- fit$deviance; rm.bic <- BIC(fit); rm.aic <- AIC(fit)
rm.stats <- cbind(rm.aic, rm.bic, rm.concord, sensitivity, specificity)
fm <- fit2$deviance; fm.bic <- BIC(fit2); fm.aic <- AIC(fit2)
fm.stats <- cbind(fm.aic, fm.bic, fm.concord, sensitivity2, specificity2)
t3 <- rbind(rm.stats, fm.stats)
rownames(t3) <- c("RM", "FM")
colnames(t3) <- c("AIC", "BIC", "Concordant", "Sensitivity", "Specificity")
```